

WEST Search History

DATE: Tuesday, September 16, 2003

<u>Set Name</u>	<u>Query</u>	<u>Hit Count</u>	<u>Set Name</u>
side by side			result set
<i>DB=USPT,PGPB,JPAB,EPAB,DWPI; PLUR=YES; OP=ADJ</i>			
L30	L29 AND human-growth-hormone	0	L30
L29	L28 AND human insulin	236	L29
L28	L27 NOT chimera	1018	L28
L27	L26 AND fusion protein	2521	L27
L26	L24 NOT Rosen-Craig.IN.	2900	L26
L25	L24 NOT Rosen-Craig.IN/	2901	L25
L24	L23 NOT Ashkenazi-Avi.IN.	2901	L24
L23	L22 NOT Rosen-Craig-A.IN.	2971	L23
L22	L17 NOT Ashkenazi-Avi-J.IN.	3165	L22
L21	L20 AND human insulin	124	L21
L20	L19 AND chimeric protein	321	L20
L19	L18 AND fusion protein	1654	L19
L18	L17 AND ((530/300 530/350)!.CCLS.)	1689	L18
L17	L16 AND L10	3281	L17
L16	L15 AND human growth hormone	5184	L16
L15	insulin	45978	L15
L14	L13 AND hGH	119	L14
L13	L12 AND human insulin	664	L13
L12	L11 AND insulin	3281	L12
L11	L10 AND human growth hormone	4467	L11
L10	(chimeric OR fusion protein OR chimera)	49651	L10
L9	L8 AND hGH	143	L9
L8	L2 AND insulin	4231	L8
L7	L2 AND insulin AND human-growth-hormone	0	L7
L6	((C07K/)!.IPC.)	8	L6
L5	(C07K)	63212	L5
L4	((C07K/5)!.CCLS.)	0	L4
L3	C07K ((5/00)!.CCLS.)	0	L3
L2	((530/300 530/350)!.CCLS.)	13606	L2
L1	(530.CCLS.)	0	L1

END OF SEARCH HISTORY

CC protein, which contains a thrombin recognition site, and human beta nerve
 CC growth factor (beta-NGF) at the C-terminal. Beta-NGF can be used to
 CC control geriatric dementia and other nervous disorders, and can be
 CC released from the fusion protein by incubation with thrombin (see
 CC AAN90577-8, AAP91034, AAP91299).

XX SQ Sequence 140 AA;
 Query Match 100.0%; Score 260; DB 10; Length 140;
 Best Local Similarity 100.0%; Pred. No. 7.3e-25;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLONP 49
 |||||
 DB 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLONP 49

RESULT 6
 AAY42861
 ID AAY42861 standard; protein: 150 AA.

XX AC AAY42861;

XX DT 19-JAN-2000 (first entry)

XX DE Chimeric protein, SEQ ID 7.

XX KW Insulin; precursor: growth hormone; chaperone: intramolecular
 KW folding; conformation: chimeric protein; cleavable; recombinant;
 KW production; yield.

XX OS Synthetic.

XX OS Homo sapiens.

XX XN WO9950302-A1.

XX XX 07-OCT-1999.

XX XX 31-MAR-1998; 98WO-CN00352.

XX PR 31-MAR-1998; 98WO-CN00352.

XX PA (TONG-) TONGHUA GANTECH BIOTECHNOLOGY LTD.

XX PI Can 2;

XX DR WPI: 1999-6i0839/52.

XX PT New chimeric proteins containing human growth hormone fragment, used
 PT particularly for the production of human insulin.

XX PS Claim 14; Page 30-31; 46pp; English.

XX CC This sequence represents a chimeric protein, which contains an
 CC N-terminal fragment of human growth hormone (hGH) of the sequence given
 CC in AAY42856, a cleavable peptide linker (AAY42857), and a human insulin
 CC precursor comprising insulin A and B chains (AAY42859). The hGH portion
 CC of the chimeric protein acts as an intramolecular chaperone (IMC) for
 CC the insulin precursor, enabling it to fold correctly. The cleavable
 CC peptide linker has a C-terminal Arg residue which enables the hGH portion
 CC of the chimeric protein to be removed after folding has taken place.
 CC Production of recombinant human insulin via an hGH-proinsulin chimera
 CC protein can provide human insulin with correctly linked cysteine bridges
 CC with fewer necessary procedural steps, and hence resulting in a higher
 CC yield of human insulin. The IMC sequences not only protect insulin
 CC sequences from intracellular degradation by a microorganism host, but
 CC also promote the folding of the fused insulin precursor, facilitate the
 CC solubility of the fusion protein and decrease the intermolecular
 CC interactions among the fusion proteins, thus allowing folding of the
 CC fused insulin precursor at commercially useful high concentrations. The
 CC procedural steps of cyanogen bromide cleavage, oxidative sulphiteolysis
 CC and related purification steps can thus be eliminated, along with the use
 CC of high concentrations of mercaptan or the use of hydrophobic absorbent

CC Resins.
 XX SQ Sequence 150 AA;

Query Match 100.0%; Score 260; DB 20; Length 150;
 Best Local Similarity 100.0%; Pred. No. 7.9e-25;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLONP 49
 |||||
 DB 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLONP 49

RESULT 7

AAP90129

ID AAP90129 standard; protein: 192 AA.

XX AC AAP90129;

XX DT 25-MAR-2003 (updated)

XX DT 06-FEB-1996 (revised)

XX DT 01-NOV-1989 (first entry)

XX XX Human growth hormone.

XX XX Human growth hormone; fusion protein; recombinant

XX KW vector.

XX OS Homo sapiens (Human).

XX XN JPO1244981-A.

XX PD 07-JUN-1989.

XX PP 02-DEC-1987; 87CP-0304937.

XX PR 02-DEC-1987; 87JP-0404937.

XX PA (WAKT) WAKUNAGA SEIYAKU KK.

XX DR WPI: 1989-209284/29.

XX DR N-PSDB; AAN90265.

XX PT Recombinant vector contg. fusion protein - consisting of human
 PT growth hormone or deriv. ligated to foreign protein, for stability
 PT and high yield.

XX PS Disclosure: Fig 1: 19pp; Japanese.

XX CC The invention consists of a vector contg. a fusion protein which is
 CC formed by ligating, downstream of a promoter, hGH or a deriv. (pref.
 CC formed by substitn. of Met-14 with Leu) and a foreign protein.
 CC Stability of the vector in the host is greatly increased so the
 CC protein yield is higher.
 CC (Updated on 25-MAR-2003 to correct PA field.)

XX SQ Sequence 192 AA;

Query Match 100.0%; Score 260; DB 10; Length 192;
 Best Local Similarity 100.0%; Pred. No. 1e-24;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLONP 49
 |||||
 DB 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLONP 49

RESULT 8

AAW92264

ID AAW92264 standard; Protein: 192 AA.

XX AC AAW92264;

XX XX

DI 08-JUN-1999 (first entry)

DE Human anti-angiogenic peptide hGH Met-1-phospho1.

XX Human: anti-angiogenic; prolactin; placental lactogen; hGH; anti-angiogenic; growth hormone; hGH; hGH-V; capillary endothelial cell proliferation; placental vascularisation; pregnancy; treatment; angiogenic disease; tumour; inhibitor; malignant; angiofibroma; arteriovenous malformation; arthritis; atherosclerotic plaques; corneal graft neovascularisation; wound healing; proliferative retinopathy; macular degeneration; trachoma; granulation; glaucoma; ocular; uveitis; fracture; Osier-Weber syndrome; psoriasis; fibroplasia; scleroderma; Kaposi's sarcoma; vascular adhesion; ulcer; leukaemia; reproductive disorder; contraceptive agent; gene therapy; pre-eclampsia; intrauterine growth retardation; placental dysfunction.

XX Homo sapiens.

OS W09851323-AL.

PN 19-NOV-1968.

XX 12-MAY-1983; 98WO-US95591.

XX 13-MAY-1997; 97US-0346394.

PR (REGC) UNIV CALIFORNIA.

XX Martini JA, Struman I, Taylor R, Weiner R1;

XX WPI: 1999-045192/04.

DR N-PSDB; AAX01706.

XX New anti-angiogenic peptides - comprise N-terminal fragments of human placental lactogen, human growth hormone, growth hormone variant or human prolactin

PT Example 3; Page 49; 87pp; English.

XX This invention describes novel human anti-angiogenic peptides derived from 10 to 150 consecutive amino acids selected from the N-terminal end of human placental lactogen (hPL), human growth hormone (hGH), growth hormone variant (hGH-V), or human prolactin. Such peptides (i) inhibit capillary endothelial cell proliferation and organisation (ii) inhibit angiogenesis in chick chorioallantoic membrane and (iii) binds to at least one specific receptor which does not bind an intact full length hGH, hPL, prolactin or hGH-V. The invention also describes a method for diagnosing a probable abnormality of placental vascularisation during pregnancy. The peptides can be used for treating an angiogenic disease in a subject, for inhibiting tumour formation or growth in a patient, for modulating vascularisation of a patient's placenta. In particular, the peptides can be used for preventing or treating e.g. malignant tumours, angiofibroma, arteriovenous malformation, arthritic such as rheumatoid arthritis, atherosclerotic plaques, corneal graft neovascularisation, delayed wound healing, proliferative retinopathy such as diabetic retinopathy, macular degeneration, granulation such as those occurring in hemophilic joints, inappropriate vascularisation in wound healing such as hypertrophic scars or keloid scars, neovascular glaucoma, ocular tumour, uveitis, non-union fractures, Osier-Weber syndrome, psoriasis, pyogenic glaucoma, retrolental fibroplasia, scleroderma, solid tumours, Kaposi's sarcoma, trachoma, vascular adhesions, chronic varicose veins, leukaemia, and reproductive disorders such as follicular and luteal cysts and choriocarcinoma. They can also be used as contraceptive agents. DNA encoding the peptides can be used in gene therapy. The measurement of abnormal levels of N-terminal fragments of hGH, hGH-V, prolactin or hPL can be used in assays for impairment of vascular development associated with pre-eclampsia, intrauterine growth retardation, and placental dysfunction.

XX Sequence 192 AA;

XX Query Match 100.0%; Score 250; DB 20; Length 192;

XX Best Local Similarity 100.0%; Pred. No. 1c-24;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPTIPSRLEFDNMLKAHLHQLAFTDYOEFEAYIPKEQKYSFLQNP 49

DB 1 MEPTIPSRLEFDNMLKAHLHQLAFTDYOEFEAYIPKEQKYSFLQNP 49

RESULT 9

AA91299

IP AAF91299 standard; protein: 261 AA.

XX AAF91299;

AC 14-DEC-1989 (first entry).

XX Human nerve growth factor and human growth hormone fusion protein.

XX Human nerve growth factor: fusion protein; thrombin;

XX geriatric dementia; nervous disorders; human growth hormone.

XX Homo sapiens (human).

XX Key Location/Qualifiers

FT Region 1..140

FT Region 141..143

FT Region 144..261

XX EP329175-A.

XX 23-AUG-1989.

XX 17-FEB-1989; 89EP-0102795.

XX 19-FEB-1988; 88JP-0035042.

XX (TOYJ) TOSOH CORP.

XX Chitsuka E;

XX WPI: 1989-243092/34.

XX New human nerve growth factor gene encoding fusion protein

PT - having cleavage site for thrombin, useful for treating geriatric dementia, etc.

XX Claim 36; page 31-32; 38pp; English.

XX Fusion protein consisting of human growth hormone at the N-terminal end (1st region), a 3 amino acid sequence representing thrombin recognition site, and human beta nerve growth factor (beta-NGF) at the C-terminal. Beta-NGF can be used to control geriatric dementia and other nervous disorders, and can be released from the fusion protein by incubation with thrombin (see AAN90577-8, AAF91034, AAF91041).

XX Sequence 261 AA;

XX Query Match 100.0%; Score 250; DB 10; Length 261;

XX Best Local Similarity 100.0%; Pred. No. 1-5e-24;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPTIPSRLEFDNMLKAHLHQLAFTDYOEFEAYIPKEQKYSFLQNP 49

DB 1 MEPTIPSRLEFDNMLKAHLHQLAFTDYOEFEAYIPKEQKYSFLQNP 49

RESULT 10

AA91033

ID AAF61033 standard; Protein: 262 AA.

XX AAF61033;

XX 25-OCT-1991 (first entry)

XX Human beta-nerve growth factor gene product.

XX Beta-NGF; E.coli; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Protein 145..262

XX JP61205465-A.

XX 11-SEP-1986.

XX 09-MAR-1985; 85JP-0045773.

XX 09-MAR-1985; 85JP-0045773.

XX (OTSU/; OTSUKA E.

XX WPI: 1986-241696/43.

XX Gene segment of human nerve growth factor - used in prodn. of NGF-producing recombinant Escherichia strain.

XX Claim 32; Page 482; 7ipp; Japanese.

XX The protein is a direct translation of the upstream tryptophan promoter-operator lacking its attenuation sequence and human beta-NGF sequence. The product may be efficiently expressed from a transformed E.coli expression system.

XX See also AAN63816-7.

XX Sequence 262 AA:

Query Match 100.0%; Score 260; DB 7; Length 262;
Best Local Similarity 100.0%; Pred. No. 1.5e-24;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFPTIPLSKLFNDNMLRAHRLHQLAFDITYQEFEEAYIPKEOKYSFLQNP 49

DB 1 MFPTIPLSKLFNDNMLRAHRLHQLAFDITYQEFEEAYIPKEOKYSFLQNP 49

RESULT 11

AAR11740

ID AAR11740 standard; Protein: 262 AA.

XX AAR11740:

XX 25-MAR-2003 (updated)

XX 25-JUN-1991 (first entry)

XX Human growth hormone/human nerve growth factor beta fusion protein.

XX hGH; hNGF; nervous system diseases; dementia.

XX Homo sapiens.

XX JP03067598-A.

XX 22-MAR-1991.

XX 07-AUG-1989; 89JP-0202835.

XX 07-AUG-1989; 89JP-0202835.

XX (TOYJ) TOSOH CORP.

XX WPI: 1991-128768/18.

XX N-PSDB; AAO11578.

XX Purificn. of human neuron growth factor beta-subunit-contg. protein -

PT by contacting with gel having cation exchange gp. in presence of urea

XX Disclosure : fig 1; 7pp; Japanese.

XX A recombinant human nerve growth factor beta subunit-contg. protein can be produced as this fusion protein. It is purified by contacting a gel having a cation exchange gp. with the fusion protein, in the presence of urea. The purified protein is useful in a medicament for treating disorders of the nervous system, eg dementia.

XX (Updated on 25-MAR-2003 to correct PF field.)

XX Sequence 262 AA:

Query Match 100.0%; Score 260; DB 12; Length 262;
Best Local Similarity 100.0%; Pred. No. 1.5e-24;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFPTIPLSKLFNDNMLRAHRLHQLAFDITYQEFEEAYIPKEOKYSFLQNP 49

DB 1 MFPTIPLSKLFNDNMLRAHRLHQLAFDITYQEFEEAYIPKEOKYSFLQNP 49

RESULT 12

AAR03255

ID AAR03255 standard; Protein: 310 AA.

XX AAR03255;

XX 19-JUL-1990 (first entry)

XX Fusion protein of B-cell stimulatory factor-2 and B-cell differentiation factor.

XX B-cell stimulatory factor-2; Interleukin-6; B-cell differentiation; Interleukin-5; fusion protein.

XX Homo sapiens.

XX JP02013375-A.

XX 01-JUL-1988; 88JP-0162556.

XX 01-JUL-1988; 88JP-0162556.

XX (TOYJ) TOSOH CORP.

XX WFI: 1990-062207/09.

XX N-PSDB; AAO02028.

XX Prepn. of human B-cell differentiation factor - from specified DNA sequence segment, by recombinant DNA technique, gives protein of specified amino acid sequence.

XX Claim 31; Page 9; 17pp; Japanese.

XX The protein is produced by fusing DNA encoding BDF (IL-) with DNA encoding BSF-2 (IL-5) and ligating the product into an expression vector

XX See also AAR05311 and AAR05313.

XX Sequence 310 AA:

Query Match 100.0%; Score 260; DB 11; Length 310;
Best Local Similarity 100.0%; Pred. No. 1.8e-24;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFPTIPLSKLFNDNMLRAHRLHQLAFDITYQEFEEAYIPKEOKYSFLQNP 49

DB 1 MFPTIPLSKLFNDNMLRAHRLHQLAFDITYQEFEEAYIPKEOKYSFLQNP 49

RESULT 13
 AAR05313
 ID AAR05313 standard; protein: 144 AA.
 XX
 AC AAR05313;
 XX
 DT 19-JUL-1990 (first entry)
 XX
 DE Segment of B-cell stimulatory factor-2 (IL-5).
 XX
 KW B-cell stimulatory factor-2; interleukin-5.
 XX
 OS Homo sapiens.
 XX
 PN JP02013375-A.
 XX
 PD 17-JAN-1990.
 XX
 PF 01-JUL-1988; 88JP-G162556.
 XX
 PR 01-JUL-1988; 88JP-0162556.
 XX
 PA (TOYO) TOSOH CORP.
 XX
 DR WPI: 1990-062207/03.
 DR N-PSDB; AAQ02C28.
 XX
 PT prepn. of human B-cell differentiation factor - from specified DNA
 PT sequence segment, by recombinant DNA technique, gives protein of
 PT specified amino acid sequence.
 XX
 PS Disclosure: Page 9; 17pp; Japanese.
 XX
 CC The sequence encoding this protein can be fused with DNA encoding B-cell
 CC differentiation factor (IL-6); and ligated into an expression vector for
 CC prodn. of a fusion protein.
 CC See also AAR05311.
 XX
 SQ Sequence 144 AA;

Query Match 98.8%; Score 257; DB 11; Length 144;
 Best Local Similarity 98.0%; Pred. No. 3-24;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFPTIPLSRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLQNP 49
 |||||
 DB 1 MFPTIPLSRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLQNP 49

RESULT 14
 ABB77327
 ID ABB77327 standard; Protein: 204 AA.
 XX
 AC ABB77327;
 XX
 DT 17-JUN-2002 (first entry)
 XX
 DE Human growth hormone with collagenase recognition site.
 XX
 KW Human; growth hormone; collagenase; recognition site.
 XX
 OS Homo sapiens.
 XX
 PN KR289691-B.
 XX
 PD 15-MAY-2001.
 XX
 PF 28-DEC-1993; 93KR-00303-B.
 XX
 PR 28-DEC-1993; 93KR-00303-B.
 XX
 PA (GLDS) LG CHEM LTD.

Query Match 98.1%; Score 255; DB 9; Length 138;
 Best Local Similarity 100.0%; Pred. No. 30-24;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FPTIPLSRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLQNP 49
 |||||

PI YOO CG, Song YH;
 XX
 DR WPI: 2002-185396/24.
 DR N-PSDB; ABL5599%.
 XX
 PT Recombinant human growth hormone having collagenase recognition region -
 XX
 PS Disclosure: Fig 3; 5pp; Korean.
 XX
 CC The invention relates to recombinant human growth hormone having a
 CC collagenase recognition region.
 XX
 SQ Sequence 204 AA;
 XX
 Query Match 98.5%; Score 256; DB 23; Length 204;
 Best Local Similarity 98.0%; Pred. No. 3-24;
 Matches 48; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFPTIPLSRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLQNP 49
 |||||
 DE 13 VFPTIPLSRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLQNP 61

RESULT 15
 AAP81226
 ID AAP81226 standard; protein: 138 AA.
 XX
 AC AAP81226;
 XX
 DT 25-MAR-2003 (updated)
 DT 20-NOV-1990 (first entry)
 XX
 DE Sequence of protein with somatomedin-like activity.
 XX
 KW Growth hormone.
 XX
 OS Synthetic.
 XX
 PN JP63167798-A.
 XX
 PD 11-JUL-1988.
 XX
 PF 29-DEC-1986; 86JP-0310177.
 XX
 PR 29-DEC-1986; 86JP-G310177.
 XX
 PA (TOYO) TOYO SODA MFG CO LTD.
 XX
 DR WPI: 1986-232632/33.
 DR N-PSDB; AAN81605.
 XX
 PT Polypeptide with somatomedin-like activity -
 PT by culturing bacterium transformed by plasmid contg. gene
 PT segment with specified DNA sequence
 XX
 PS Claim 2(1): Page 609; 9pp; Japanese.
 XX
 CC The polypeptide (AAP81226) with somatomedin-like activity and the DNA
 CC (AAN81605) encoding it are claimed. A vet. residual gp. may be added to
 CC the N-terminal. The polypeptide acts on the bone structure of mammals,
 CC including humans, to promote bone growth. The polypeptide has high
 CC production rate and is easily extracted from bacterial culture medium
 CC and refined for use as a bone growth accelerator.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 138 AA;

Query Match 98.1%; Score 255; DB 9; Length 138;
 Best Local Similarity 100.0%; Pred. No. 30-24;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FPTIPLSRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLQNP 49
 |||||

DB 1 FTIPLSRLEFDNAMLPAHRLHQLEAFDTYCEFFERAYCPKQKYSFLQNP 43

Search completed: September 15, 2003, 12:00:56
Job time : 35.4229 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Inc.

OM protein - protein search, using sw model

Run on: September 15, 2003, 11:57:43 : Search time 11,2401 seconds
(without alignments)
184,449 Million cell updates/sec

Title: US-09-423-100-2

Perfect score: 280
Sequence: 1 MFPTPLSRFLDNPMLRAHR.....QEFEDAYPKCKYKSFQNP 49

Scoring table: BLCSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310658 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2003000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/iaa/5A.COMB pep: *
2: /cgn2_6/ptodata/1/iaa/5B.COMB pep: *
3: /cgn2_6/ptodata/1/iaa/6A.COMB pep: *
4: /cgn2_6/ptodata/1/iaa/6B.COMB pep: *
5: /cgn2_6/ptodata/1/iaa/6C.COMB pep: *
6: /cgn2_6/ptodata/1/iaa/6D.COMB pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the results being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	260	100.0	192	1	US-08-093-363-1
2	255	98.1	191	4	US-09-284-878-5
3	255	98.1	194	2	US-08-383-621-4
4	255	98.1	194	3	US-08-459-956-4
5	255	98.1	217	3	US-06-589-023-10
6	255	98.1	217	3	US-08-784-582-10
7	255	98.1	217	3	US-08-785-271-19
8	255	98.1	217	3	US-08-759-528-11
9	255	98.1	217	4	US-09-284-878-1
10	255	98.1	241	4	US-09-424-620B-25
11	255	98.1	245	4	US-09-285-030-56
12	255	98.1	274	3	US-08-784-582-71
13	255	98.1	360	3	US-08-784-562-73
14	249	95.8	193	4	US-05-465-461-1
15	249	95.8	217	1	US-08-187-756C-4
16	249	95.8	217	1	US-08-469-486-5
17	249	95.8	217	2	US-08-469-658-51
18	249	95.8	217	2	US-08-710-324A-4
19	249	95.8	217	4	US-09-411-657-4
20	248	95.4	191	3	US-08-800-215C-16
21	248	95.4	191	3	US-08-800-215C-18
22	248	95.4	191	3	US-08-800-215C-20
23	248	95.4	400	4	US-09-420-819-37
24	248	95.4	401	4	US-09-420-819-36
25	237	91.2	71	1	US-08-314-586-24
26	233	89.6	70	1	US-07-920-519-24
27	233	89.6	70	3	US-08-115-753-26

28 164.5 63.3 191 1 US-08-458-824-8
29 164 63.1 176 3 US-08-791-728-1
30 164 63.1 176 4 US-08-990-774-1
31 161.5 62.3 192 1 US-07-963-331D-4
32 159.5 61.3 190 1 US-08-388-267C-2
33 159.5 61.3 190 4 US-09-277-720-2
34 159.5 61.3 191 6 5210180-1
35 159.5 61.3 193 1 US-07-621-197C-2
36 159.5 61.3 193 1 US-08-363-982-2
37 159.5 61.3 193 2 US-08-383-621-1
38 159.5 61.3 193 3 US-08-459-906-1
39 159.5 61.3 216 2 US-09-105-651-1
40 159.5 61.3 216 2 US-09-105-651-3
41 158 60.8 176 3 US-08-791-728-2
42 158 60.8 176 4 US-08-990-774-2
43 158 60.8 177 1 US-08-187-756C-6
44 158 60.8 177 2 US-08-710-324A-6
45 158 60.8 177 4 US-09-411-657-6

ALIGNMENTS

RESULT 1
US-08-093-363-1
: Sequence 1, Application US/08093383
: Patent No. 5489529
: GENERAL INFORMATION:
: APPLICANT: DeBoer, Herman A.
: APPLICANT: Heyneker, Herbert L.
: APPLICANT: Seeburg, Peter H.
: TITLE OF INVENTION: DNA for Expression of Bovine Growth Hormone
: NUMBER OF SEQUENCES: 35
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Genentech, Inc.
: STREET: 460 Point San Bruno Blvd
: CITY: South San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94080
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: patin (Genentech)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/093.383
: FILING DATE: 14-JUL-1993
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/519827
: FILING DATE: 28-NOV-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/198824
: FILING DATE: 05-APR-1988
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 06/632361
: FILING DATE: 19-JUL-1984
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 06/303687
: FILING DATE: 18-SEP-1981
: ATTORNEY/AGENT INFORMATION:
: NAME: Johnston, Sean A.
: REGISTRATION NUMBER: P35,910
: REFERENCE/DOCKET NUMBER: 46C4
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415/225-1552
: TELEFAX: 415/952-9881
: TELEX: 910/371-7166
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 192 amino acids
: TYPE: amino acid

Sequence 8, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 4, Appli
Sequence 2, Appli
Sequence 2, Appli
Patent No. 5210180
Sequence 2, Appli
Sequence 2, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 2, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 6, Appli


```

Query Match      98.1%; Score 255; DB 4; Length 217;
Best Local Similarity 100.0%; Pred. No. 2.9e-29;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FPTPLSRFDNMLRAHRLHQLAFDTYQFEFAYIPKEOKYSFLQNP 49
      |||||.....|||||.....|||||.....|||||.....
DB 4 FPTPLSRFDNMLRAHRLHQLAFDTYQFEFAYIPKEOKYSFLQNP 51

RESULT 5
US-08-589-028-10
; Sequence 10, Application US/08589028
; Patent No. 6087129
; GENERAL INFORMATION:
; APPLICANT: Newgard, Christopher B.
; APPLICANT: Halban, Philippe
; APPLICANT: No. 6087129mington, Karl D.
; APPLICANT: Clark, Samuel A.
; APPLICANT: Thigpen, Anice E.
; APPLICANT: Quade, Christian
; APPLICANT: Kruse, Fred
; TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P. O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/03/589,028
; FILING DATE: Concurrently/ Herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 47,542
; REFERENCE/DOCKET NUMBER: UTSD:426\HYL
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-589-028-10

Query Match      98.1%; Score 255; DB 4; Length 217;
Best Local Similarity 100.0%; Pred. No. 2.9e-29;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FPTPLSRFDNMLRAHRLHQLAFDTYQFEFAYIPKEOKYSFLQNP 49
      |||||.....|||||.....|||||.....|||||.....
DB 27 FPTPLSRFDNMLRAHRLHQLAFDTYQFEFAYIPKEOKYSFLQNP 74

RESULT 6
US-08-784-582-10
; Sequence 10, Application US/06:84582
; Patent No. 6110707
; GENERAL INFORMATION:
; APPLICANT: Newgard, Christopher B.
; APPLICANT: Halban, Philippe A.
; APPLICANT: No. 6110707mington, Karl D.

```

```

; APPLICANT: Clark, Samuel A.
; APPLICANT: Thigpen, Anice E.
; APPLICANT: Quade, Christian
; APPLICANT: Kruse, Fred
; APPLICANT: McGarry, Dennis
; TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P. O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/784,582
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/028,427
; FILING DATE: 15-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/589,028
; FILING DATE: 19-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: UTSD:514
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-784-582-10

Query Match      98.1%; Score 255; DB 3; Length 217;
Best Local Similarity 100.0%; Pred. No. 2.9e-29;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FPTPLSRFDNMLRAHRLHQLAFDTYQFEFAYIPKEOKYSFLQNP 49
      |||||.....|||||.....|||||.....|||||.....
DB 27 FPTPLSRFDNMLRAHRLHQLAFDTYQFEFAYIPKEOKYSFLQNP 74

RESULT 7
US-08-785-271-10
; Sequence 10, Application US/08785271
; Patent No. 6194176
; GENERAL INFORMATION:
; APPLICANT: Newgard, Christopher B.
; APPLICANT: Halban, Philippe A.
; APPLICANT: No. 6194176mington, Karl D.
; APPLICANT: Clark, Samuel A.
; APPLICANT: Thigpen, Anice E.
; APPLICANT: Quade, Christian
; APPLICANT: Kruse, Fred
; TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P. O. Box 4433
; CITY: Houston

```

```
1 STATE: Texas
2 COUNTRY: USA
3 ZIP: 77210
4 COMPUTER READABLE FORM:
5 MEDIUM TYPE: Floppy disk
6 COMPUTER: IBM PC compatible
7 OPERATING SYSTEM: PC-DOS/MS-DOS
8 SOFTWARE: Patent In Release #1.0, Version #1.0
9 CURRENT APPLICATION DATA:
10 APPL. APPLICATION NUMBER: US/08/785-271
11 FILING DATE: Concurrently Hologram
12 CLASSIFICATION: 435
13 PRIOR APPLICATION DATA:
14 APPLICATION NUMBER: US 08/589,028
15 FILING DATE: 19-JAN-1996
16 ATTORNEY/AGENT INFORMATION:
17 NAME: Highlander, Steven L.
18 REGISTRATION NUMBER: 37,542
19 REFERENCE/DOCKET NUMBER: USH:513
20 TELECOMMUNICATION INFORMATION:
21 TELEPHONE: 512/474-7577
22 TELEFAX: 512/474-7577
23 INFORMATION FOR SEQ ID NO: 10:
24 SEQUENCE CHARACTERISTICS:
25 LENGTH: 217 amino acids
26 TYPE: amino acid
27 STRANDEDNESS: linear
28 TOPOLOGY: linear
29 US-08-785-271-10

Query Match 98.1%; Score 255; DH 3; Length 217;
Best Local Similarity 100.0%; Pred. No. 2,9e-29;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FPIPLSRFDNAMLRAHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNP 49
DB 27 FPIPLSRFDNAMLRAHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNP 74

RESULT 8
US-08-759-628-11
: Sequence 1, Application US/08/59628
: Patent No. 5225446
: GENERAL INFORMATION:
: APPLICANT: Altmann, Scott W.
: APPLICANT: Rock, Fernando L.
: APPLICANT: Bazan, J. Fernando
: APPLICANT: Kastlein, Robert A.
: TITLE OF INVENTION: MUTATIONAL VARIANTS OF MAMMALIAN PROTEINS
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: DNAX Research Institute
: STREET: 901 California Avenue
: CITY: Palo Alto
: STATE: California
: COUNTRY: USA
: ZIP: 94304-1104
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/759-628
: FILING DATE: 05-DEC-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 56/606,574
: FILING DATE: 06-DEC-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Chang, Edwin P.
: REGISTRATION NUMBER: 34,690
: REFERENCE/DOCKET NUMBER: DX05520
```

```
1 TELECOMMUNICATION INFORMATION:
2 TELEPHONE: 415-852-9196
3 TELEFAX: 415-496-1200
4 INFORMATION FOR SEQ ID NO: 11:
5 SEQUENCE CHARACTERISTICS:
6 LENGTH: 217 amino acids
7 TYPE: amino acid
8 STRANDEDNESS: single
9 TOPOLOGY: linear
10 MOLECULE TYPE: Protein
11 FEATURE:
12 NAME/KEY: Peptide
13 LOCATION: 32..53
14 FEATURE:
15 NAME/KEY: Peptide
16 LOCATION: 94..115
17 FEATURE:
18 NAME/KEY: Peptide
19 LOCATION: 133..153
20 FEATURE:
21 NAME/KEY: Peptide
22 LOCATION: 192..210
23 OTHER INFORMATION: /color "The peptides above are
24 OTHER INFORMATION: depicted in Figure 1"
25 US-09-423-100-1-11

Query Match 98.1%; Score 255; DS 3; Length 217;
Best Local Similarity 100.0%; Pred. No. 2,9e-29;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FPIPLSRFDNAMLRAHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNP 49
DB 27 FPIPLSRFDNAMLRAHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNP 74

RESULT 9
US-09-284-878-1
: Sequence 1, Application US/09284878
: Patent No. 5342375
: GENERAL INFORMATION:
: APPLICANT: Olazaran, Martha Guerrero
: APPLICANT: Saldana, Hugo Barrera
: APPLICANT: Salgado, Jose Maria Viader
: TITLE OF INVENTION: Genetically Modified Methylothrophic P. pastoris Yeast for t
: TITLE OF INVENTION: Production and Secretion of the Human Growth Hormone
: FILE REFERENCE: 1829,00:0000
: CURRENT APPLICATION NUMBER: US/09/284,878
: CURRENT FILING DATE: 1999-07-21
: PRIOR APPLICATION NUMBER: PCT/MX97/00033
: PRIOR FILING DATE: 1997-10-24
: NUMBER OF SEQ ID NOS: 9
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 1
: LENGTH: 217
: TYPE: PPT
: ORGANISM: Homo sapiens
: US-09-284-878-1

Query Match 96.1%; Score 255; DS 4; Length 217;
Best Local Similarity 100.0%; Pred. No. 2,9e-29;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FPIPLSRFDNAMLRAHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNP 49
DB 27 FPIPLSRFDNAMLRAHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNP 74

RESULT 10
US-09-424-620B-25
: Sequence 25, Application US/09424620B
: Patent No. 639,585
: GENERAL INFORMATION:
: APPLICANT: HANIL SYNTHETIC FIBER CO., LTD.
```


APPLICANT: Newgard, Christopher B.
APPLICANT: Halban, Philippe A.
APPLICANT: No. 6110707;ington, Karl D.
APPLICANT: Clark, Samuel A.
APPLICANT: Thiepen, Anice E.
APPLICANT: Ouade, Christian
APPLICANT: Kruse, Fred
APPLICANT: McGarity, Dennis
TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
NUMBER OF SEQUENCES: 79
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/784,582
FILING DATE: Concurrently Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/026,427
FILING DATE: 15-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/589,028
FILING DATE: 19-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: UTSD:514
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 73:
SEQUENCE CHARACTERISTICS:
LENGTH: 366 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-784-582-73

Query Match 98.18; Score 255; DB 3; Length 360;
Best Local Similarity 100.0%; Pred. No. 5.4e-29;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FPIPLSRLEFDNMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLONP 49
DB 27 FPIPLSRLEFDNMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLONP 74

RESULT 14
US-09-465-461-1
Sequence 1; Application US/09465461
Patent No. 634844
GENERAL INFORMATION:
APPLICANT: CHAPPEL, Scott
TITLE OF INVENTION: Human Growth Hormone to stimulate hematopoiesis and enhance recovery
FILE REFERENCE: after hematopoietic stem cell transplantation in humans
CURRENT APPLICATION NUMBER: 05/09/465,461
CURRENT FILING DATE: 1993-12-17
PRIOR APPLICATION NUMBER: 60/112,668
PRIOR FILING DATE: 1998-12-17
NUMBER OF SEQ ID NOS: 1
SOFTWARE: Patent in version 3.1
SEQ ID NO 1

LENGTH: 191;
TYPE: PRI
ORIGINISM: homo sapiens
US-09-465-461-1
Query Match 95.8%; Score 249; DB 4; Length 191;
Best Local Similarity 97.9%; Pred. No. 1.8e-28;
Matches 47; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 FPIPLSRLEFDNMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLONP 49
DB 27 FPIPLSRLEFDNMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLONP 74

RESULT 15
US-08-187-756C-4
Sequence 4; Application US/08187756C
Patent No. 5597709
GENERAL INFORMATION:
APPLICANT: ROSEN, El AL.
TITLE OF INVENTION: Human Growth Hormone
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNEL BAIN, GILFILLIAN,
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/187,756C
FILING DATE: January 27, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: FERRARO, GREGORY D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325600-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 217 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-187-756C-4

Query Match 95.8%; Score 249; DB 1; Length 217;
Best Local Similarity 97.9%; Pred. No. 2.1e-28;
Matches 47; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FPIPLSRLEFDNMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLONP 49
DB 27 FPIPLSRLEFDNMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLONP 74

Search completed: September 15, 2003, 12:05:30
Job time : 12.2401 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Computer Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2003, 12:33:35 : Search time 20.544 seconds
(without alignments)
347.945 Million cell updates/sec

Title: US-09-423-100-1

Perfect score: 260

Sequence: 1 MFPIPLSRFDNMLRAHR.....ORRENAVLPKCKYFLQNP 49

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 541936 seqs, 1459.2426 residues

Total number of hits satisfying chosen parameters: 541936

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:

1: /cgn2_6/ptocata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptocata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptocata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptocata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptocata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptocata/1/pubpaa/PC_US_PUBCOMB.pep.*
7: /cgn2_6/ptocata/1/pubpaa/US06_NEW_PUB.pep.*
8: /cgn2_6/ptocata/1/pubpaa/US06_PUBCOMB.pep.*
9: /cgn2_6/ptocata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptocata/1/pubpaa/US04_PUBCOMB.pep.*
11: /cgn2_6/ptocata/1/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/ptocata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptocata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptocata/1/pubpaa/US10A_PUBCOMB.pep.*
15: /cgn2_6/ptocata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptocata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptocata/1/pubpaa/US06_NEW_PUB.pep.*
18: /cgn2_6/ptocata/1/pubpaa/US06_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	260	100.0	49	US-10-054-673-1	Sequence 1, Appl
2	260	100.0	92	US-10-054-673-2	Sequence 2, Appl
3	260	100.0	107	US-10-054-673-6	Sequence 6, Appl
4	260	100.0	150	US-10-054-673-7	Sequence 7, Appl
5	255	98.1	191	US-09-984-010-23	Sequence 23, Appl
6	255	98.1	191	US-10-153-207-1	Sequence 1, Appl
7	255	98.1	191	US-10-400-377-1	Sequence 1, Appl
8	255	98.1	191	US-10-400-768-1	Sequence 1, Appl
9	255	98.1	214	US-10-153-207-6	Sequence 6, Appl
10	255	98.1	217	US-09-929-918-9	Sequence 9, Appl
11	255	98.1	245	US-09-280-030-05	Sequence 66, Appl
12	249	95.8	217	US-09-853-688-2	Sequence 2, Appl
13	249	95.8	217	US-09-969-7480-4	Sequence 4, Appl
14	248	95.4	217	US-09-804-409A-16	Sequence 16, Appl
15	242	93.1	217	US-09-853-688-4	Sequence 4, Appl

16	239	92.9	191	12	US-09-824-200-12	Sequence 12, Appl
17	199	76.5	217	9	US-09-850-887-3	Sequence 3, Appl
18	197	75.8	191	15	US-10-043-487-150	Sequence 350, Appl
19	197	75.8	191	12	US-10-153-207-2	Sequence 2, Appl
20	197	75.8	229	15	US-10-183-313-411	Sequence 411, Appl
21	166	72.3	246	15	US-10-188-246-18	Sequence 18, Appl
22	159.5	61.3	190	12	US-10-153-207-3	Sequence 3, Appl
23	142	54.6	54	9	US-09-876-478-14	Sequence 14, Appl
24	114	43.8	54	9	US-09-876-478-2	Sequence 2, Appl
25	114	43.8	54	9	US-09-876-478-3	Sequence 3, Appl
26	106	40.6	28	9	US-09-876-478-4	Sequence 4, Appl
27	106	40.6	28	9	US-09-876-478-7	Sequence 7, Appl
28	106	40.6	54	9	US-09-876-478-6	Sequence 6, Appl
29	88	33.8	41	15	US-10-191-879-22	Sequence 22, Appl
30	88	33.8	188	15	US-10-191-879-19	Sequence 19, Appl
31	88	33.8	210	15	US-10-191-879-10	Sequence 10, Appl
32	87	33.5	366	10	US-09-887-569A-2	Sequence 2, Appl
33	87	33.5	388	12	US-10-322-746-4	Sequence 4, Appl
34	83	31.9	56	9	US-09-876-478-15	Sequence 15, Appl
35	80	30.8	198	15	US-10-140-293-32	Sequence 32, Appl
36	80	30.8	199	15	US-10-140-293-33	Sequence 33, Appl
37	73	28.1	46	15	US-10-100-679-61	Sequence 61, Appl
38	73	28.1	46	15	US-10-100-679-62	Sequence 62, Appl
39	73	28.1	46	15	US-10-100-679-63	Sequence 63, Appl
40	72.5	27.9	125	14	US-10-036-869-25	Sequence 25, Appl
41	72.5	27.9	253	14	US-10-036-869-27	Sequence 27, Appl
42	71	27.3	197	15	US-10-140-293-13	Sequence 13, Appl
43	70	26.9	199	15	US-10-140-293-22	Sequence 22, Appl
44	70	26.9	199	15	US-10-140-293-23	Sequence 23, Appl
45	70	26.9	199	15	US-10-140-293-25	Sequence 25, Appl

ALIGNMENTS

RESULT 1
US-10-054-673-1
Sequence 1, Application US/10054873
Publication No. US26020164712A1
GENERAL INFORMATION:
APPLICANT: Gao, Zhong Ru
TITLE OF INVENTION: Intramolecular Protein Containing an
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentia Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/054,873
FILING DATE: 22-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/CN98/00052
FILING DATE: 31-MAR-1998
APPLICATION NUMBER: US 09/423,100
FILING DATE: 11-DEC-2000
ATTORNEY/AGENT INFORMATION:
NAME: Mycroft, Frank J
REGISTRATION NUMBER: 46,946
REFERENCE/DOCKET NUMBER: 020167-000130US
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid


```
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: WO PCT/CN98/00052
? FILING DATE: 31-MAR-1998
? APPLICATION NUMBER: US 09/423,100
? FILING DATE: 11-DEC-2000
? ATTORNEY/AGENT INFORMATION:
? NAME: Mycright, Frank J
? REGISTRATION NUMBER: 46,946
? REFERENCE/DOCKET NUMBER: 020167-0001A005
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 150 amino acids
? TYPE: amino acid
? STRANDEDNESS: <Unknown>
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-054-873-7

Query Match 100.0%; Score 255; DB 11; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.3e-26;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 MEPTPLSR:FDNAMLRAHLRLHQLAFDTYOEFEAYIPKQKYSFLQNP 49
      ||||| ||||| || ||||| ||||| ||||| ||||| |||||
DB 1 MEPTPLSR:FDNAMLRAHLRLHQLAFDTYOEFEAYIPKQKYSFLQNP 49

RESULT 5
US-09-984-010-23
? Sequence 23, Application US/09984010
? Publication No. US20030164578A1
? GENERAL INFORMATION:
? APPLICANT: Ballance, David James
? TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
      AND SERUM ALBUMIN
? NUMBER OF SEQUENCES: 26
? CORRESPONDENCE ADDRESS:
? ADDRESSER: FINNEGAN, HENDERSON, FARROW, GARRETT & DUNN, LLP
? STREET: 1300 I Street, NW
? CITY: Washington
? STATE: DC
? COUNTRY: USA
? ZIP: 20005-3314
? COMPUTER READABLE FORM:
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent; Release #1.0, Version #1.30 (EPO)
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/984,010
? FILING DATE: 21-MAY-2002
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 09/091,873
? FILING DATE: 25-JUN-1998
? APPLICATION NUMBER: PCT/GB94/04164
? FILING DATE: 19-DEC-1996
? INFORMATION FOR SEQ ID NO: 23:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 191 amino acids
? TYPE: amino acid
? STRANDEDNESS: <Unknown>
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? HYPOTHETICAL: NO
? ANTI-SENSE: NO
? SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-984-010-23

Query Match 98.1%; Score 255; DB 11; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.3e-26;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

```
QY 2 FETIPLSR:FDNAMLRAHLRLHQLAFDTYOEFEAYIPKQKYSFLQNP 49
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 FETIPLSR:FDNAMLRAHLRLHQLAFDTYOEFEAYIPKQKYSFLQNP 48

RESULT 6
US-10-153-207-1
? Sequence 1, Application US/10153207
? Publication No. US20030153003A1
? GENERAL INFORMATION:
? APPLICANT: James A. Wells
? APPLICANT: Brian C. Cunningham
? TITLE OF INVENTION: GROWTH HORMONE VARIANTS
? FILE REFERENCE: 669.12-US-C7
? CURRENT APPLICATION NUMBER: US/10/153,207
? CURRENT FILING DATE: 2002-05-22
? PRIOR APPLICATION NUMBER: 08/479,894
? PRIOR FILING DATE: 1995-06-07
? PRIOR APPLICATION NUMBER: 08/190,724
? PRIOR FILING DATE: 1994-02-02
? PRIOR APPLICATION NUMBER: 07/960,227
? PRIOR FILING DATE: 1992-10-13
? PRIOR APPLICATION NUMBER: 07/875,204
? PRIOR FILING DATE: 1992-04-27
? PRIOR APPLICATION NUMBER: 07/428,066
? PRIOR FILING DATE: 1989-10-26
? PRIOR APPLICATION NUMBER: 07/264,611
? PRIOR FILING DATE: 1988-10-28
? NUMBER OF SEQ ID NOS: 20
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 1
? LENGTH: 191
? TYPE: PRT
? ORGANISM: Homo Sapiens
? ORGANISM: Homo Sapiens
US-10-153-207-1

Query Match 95.1%; Score 255; DB 12; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.3e-26;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0:

QY 2 FETIPLSR:FDNAMLRAHLRLHQLAFDTYOEFEAYIPKQKYSFLQNP 49
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 FETIPLSR:FDNAMLRAHLRLHQLAFDTYOEFEAYIPKQKYSFLQNP 48

RESULT 7
US-10-400-377-1
? Sequence 1, Application US/10400377
? Publication No. US20030162949A1
? GENERAL INFORMATION:
? APPLICANT: Cox 171, George N
? APPLICANT: Boldor Biotechnology, Inc.
? TITLE OF INVENTION: Derivatives of Growth Hormone and Related Proteins
? FILE REFERENCE: 4152-1-PUS
? CURRENT APPLICATION NUMBER: US/10/400,377
? CURRENT FILING DATE: 2003-03-26
? PRIOR APPLICATION NUMBER: US/09/462,941
? PRIOR FILING DATE: 2000-01-14
? PRIOR APPLICATION NUMBER: 60/052,516
? PRIOR FILING DATE: 1997-07-14
? NUMBER OF SEQ ID NOS: 41
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 1
? LENGTH: 191
? TYPE: PRT
? ORGANISM: Homo sapiens
? ORGANISM: Homo sapiens
US-10-400-377-1

Query Match 98.1%; Score 255; DB 12; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.3e-26;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0:

QY 2 FETIPLSR:FDNAMLRAHLRLHQLAFDTYOEFEAYIPKQKYSFLQNP 49
```

Db 1 FPTIPSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
|||||

RESULT 8

US-10-400-708-1
: Sequence 5, Application US/10400729
: Publication No. US20030156855A1
: GENERAL INFORMATION:
: APPLICANT: Cox III, George N
: APPLICANT: Bolder Biotechnology, Inc.
: TITLE OF INVENTION: Derivatives of Growth Hormone and Related Proteins
: FILE REFERENCE: 4152-1-PUS
: CURRENT APPLICATION NUMBER: US/10/400,728
: CURRENT FILING DATE: 2003-03-26
: PRIOR APPLICATION NUMBER: US/09/452,941
: PRIOR FILING DATE: 2000-01-14
: PRIOR APPLICATION NUMBER: 66/052,516
: PRIOR FILING DATE: 1997-07-14
: NUMBER OF SEQ ID NOS: 41
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 191
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-400-708-1

Query Match 98.1% Score 255; DB 12: Length 191;
Best Local Similarity 100.0%; Pred. No. 1.3e-26;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FPTIPSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
|||||

DB 1 FPTIPSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 48
|||||

RESULT 9

US-10-153-207-6
: Sequence 6, Application US/10153207
: Publication No. US20030153003A1
: GENERAL INFORMATION:
: APPLICANT: James A. Wells
: APPLICANT: Brian C. Cunningham
: TITLE OF INVENTION: GROWTH HORMONE VARIANTS
: FILE REFERENCE: 569-12-US-07
: CURRENT APPLICATION NUMBER: US/10/153,207
: CURRENT FILING DATE: 2002-05-22
: PRIOR APPLICATION NUMBER: 06/479,864
: PRIOR FILING DATE: 1995-06-07
: PRIOR APPLICATION NUMBER: 06/190,723
: PRIOR FILING DATE: 1994-02-02
: PRIOR APPLICATION NUMBER: 07/960,427
: PRIOR FILING DATE: 1992-10-13
: PRIOR APPLICATION NUMBER: 07/875,204
: PRIOR FILING DATE: 1992-04-27
: PRIOR APPLICATION NUMBER: 07/428,066
: PRIOR FILING DATE: 1989-10-26
: PRIOR APPLICATION NUMBER: 07/264,611
: PRIOR FILING DATE: 1988-10-28
: NUMBER OF SEQ ID NOS: 20
: SOFTWARE: PatentIn Ver. 4.0
: SEQ ID NO 6
: LENGTH: 214
: TYPE: PRT
: ORGANISM: Homo Sapiens
US-10-153-207-6

Query Match 98.1% Score 255; DB 12: Length 214;
Best Local Similarity 100.0%; Pred. No. 1.5e-26;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FPTIPSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
|||||

Db 24 FPTIPSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 71

RESULT 10

US-09-929-918-9
: Sequence 9, Application US/09929918
: Patent No. US20020090678A1
: GENERAL INFORMATION:
: APPLICANT: Kordyum, Vitaliy A.
: APPLICANT: Chernykh, Svitlana I.
: APPLICANT: Slavchenko, Iryna Yu.
: APPLICANT: Vozianov, Olexsandr
: TITLE OF INVENTION: PHAGE-DEPENDENT SUPER PRODUCTION OF
: FILE REFERENCE: PHAGE 006A
: CURRENT APPLICATION NUMBER: US/09/929,918
: CURRENT FILING DATE: 2001-08-15
: PRIOR APPLICATION NUMBER: 09/318,288
: PRIOR FILING DATE: 1999-05-25
: NUMBER OF SEQ ID NOS: 11
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 9
: LENGTH: 217
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-929-918-9

Query Match 98.1% Score 255; DB 9: Length 217;
Best Local Similarity 100.0%; Pred. No. 1.5e-26;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FPTIPSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
|||||

DB 27 FPTIPSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 74
|||||

RESULT 11

US-09-280-030-66
: Sequence 66, Application US/09280030A
: Patent No. US20010021515A1
: GENERAL INFORMATION:
: APPLICANT: Sato, Seiji
: APPLICANT: Higashikuni, Naoniko
: APPLICANT: Kudo, Toshiyuki
: APPLICANT: Kondo, Masaaki
: TITLE OF INVENTION: DNAs ENCODING NEW FUSION PROTEINS AND PROCESSES FOR
: PREPARING USEFUL POLYPEPTIDES THROUGH EXPRESSION OF THE
: TITLE OF INVENTION: DNAs
: FILE REFERENCE: 382.1026
: CURRENT APPLICATION NUMBER: US/09/280,030A
: CURRENT FILING DATE: 1999-03-26
: EARLIER APPLICATION NUMBER: JP10-87339/1998
: EARLIER FILING DATE: 1998-03-31
: NUMBER OF SEQ ID NOS: 66
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 66
: LENGTH: 245
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: Designated as
: OTHER INFORMATION: an amino acid sequence of MMTsp-MWmp20-TEV-G-GH
US-09-280-030-66

Query Match 98.1% Score 255; DB 9: Length 245;
Best Local Similarity 100.0%; Pred. No. 1.7e-26;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FPTIPSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
|||||

DB 55 FPTIPSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 102
|||||

```

RESULT 12
US-09-853-688-2
; Sequence 2, Application US/09t51688
; Patent No. US20020081605A1
; GENERAL INFORMATION:
; APPLICANT: COOPER, DAVID N.
; APPLICANT: PROCTER, ANNIE M.
; APPLICANT: GREGORY, JOHN
; APPLICANT: MILLAR, DAVID S.
; TITLE OF INVENTION: METHOD FOR DETECTING GROWTH HORMONE VARIATIONS IN
; FILE REFERENCE: WCM78
; CURRENT APPLICATION NUMBER: US/05/853,688
; CURRENT FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-853-688-2

Query Match      95.8%   Score 243; DB 9; Length 217;
Best Local Similarity 97.9%   Pred. No. 9.8e-26;
Matches 47; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY  2 FPTPLSLRFDNAMLRAHRLHQLAFDTYQEEFEAYIPKEOKYSFLQNP 49
    |||||.....|.....|.....|.....|.....|.....|
Db  27 FPTPLSLRFDNAMLRAHRLHQLAFDTYQEEFEAYIPKEOKYSFLQNP 74

RESULT 13
US-09-969-748C-4
; Sequence 4, Application US/09969748C
; Publication No. US20030161809A1
; GENERAL INFORMATION:
; APPLICANT: ARIZEXE PHARMACEUTICALS, INC.
; APPLICANT: HOUSTON, LOU L.
; APPLICANT: SHERIDAN, Philip J.
; APPLICANT: HAWLEY, Stephen
; APPLICANT: GUINN, Jacqueline, M.
; APPLICANT: CHAPIN, Steven
; APPLICANT: BASU, Anandash
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF BIOLOGICALLY ACTIVE
; FILE REFERENCE: 057220-0303
; CURRENT APPLICATION NUMBER: US/09/969,748C
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/267,501
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/248,819
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: US 60/248,478
; PRIOR FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: US 60/237,929
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 4
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-969-748C-4

Query Match      95.8%   Score 240; DB 12; Length 217;
Best Local Similarity 97.9%   Pred. No. 9.8e-26;
Matches 47; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY  2 FPTPLSLRFDNAMLRAHRLHQLAFDTYQEEFEAYIPKEOKYSFLQNP 49
    |||||.....|.....|.....|.....|.....|.....|
Db  27 FPTPLSLRFDNAMLRAHRLHQLAFDTYQEEFEAYIPKEOKYSFLQNP 74

```

```

RESULT 14
US-09-804-409A-16
; Sequence 16, Application US/09804409A
; Patent No. US20020155100A1
; GENERAL INFORMATION:
; APPLICANT: KIEFFER, TIMOTHY J.
; APPLICANT: CHEUNG, ANTHONY T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATED PROTEIN
; FILE REFERENCE: EXPRESSION IN GUT
; CURRENT APPLICATION NUMBER: US/09/804,409A
; CURRENT FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 16
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-804-409A-16

```

```

Query Match      95.4%   Score 248; DB 10; Length 217;
Best Local Similarity 97.9%   Pred. No. 1.3e-25;
Matches 47; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY  2 FPTPLSLRFDNAMLRAHRLHQLAFDTYQEEFEAYIPKEOKYSFLQNP 49
    |||||.....|.....|.....|.....|.....|.....|
Db  27 FPTPLSLRFDNAMLRAHRLHQLAFDTYQEEFEAYIPKEOKYSFLQNP 74

RESULT 15
US-09-853-688-4
; Sequence 4, Application US/09853688
; Patent No. US20020081605A1
; GENERAL INFORMATION:
; APPLICANT: COOPER, DAVID N.
; APPLICANT: PROCTER, ANNIE M.
; APPLICANT: GREGORY, JOHN
; APPLICANT: MILLAR, DAVID S.
; TITLE OF INVENTION: METHOD FOR DETECTING GROWTH HORMONE VARIATIONS IN
; FILE REFERENCE: WCM78
; CURRENT APPLICATION NUMBER: US/09/853,688
; CURRENT FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-853-688-4

```

```

Query Match      93.1%   Score 242; DB 9; Length 217;
Best Local Similarity 95.8%   Pred. No. 8.6e-25;
Matches 46; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY  2 FPTPLSLRFDNAMLRAHRLHQLAFDTYQEEFEAYIPKEOKYSFLQNP 49
    |||||.....|.....|.....|.....|.....|.....|
Db  27 FPTPLSLRFDNAMLRAHRLHQLAFDTYQEEFEAYIPKEOKYSFLQNP 74

```

Search completed: September 15, 2003, 12:23:35
 Job time : 21.5484 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2003, 11:56:45 : Search file 6.57865 sequences
(without alignments)
583.744 Million cell updates/sec

Title: US-09-423-100-1

Perfect score: 260

Sequence: 1 MEFTTSLNLFNMLRAHR.....OFFEPAVLPKQKSYSLQNP 49

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9618682 residues

Total number of hits satisfying chosen parameters: 283508

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match: 04

Maximum Match: 100%

Display first 45 summaries

Database :

PIR76.*

1: p1r1.*

2: p1r2.*

3: p1r3.*

4: p1r4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	255	98.1	217	1	STBO
2	255	98.1	217	1	STSH
3	228	87.7	217	1	STSH
4	228	87.7	217	1	STSH
5	213	81.9	212	2	STSH
6	213	81.9	212	2	STSH
7	205	78.8	217	2	STSH
8	203	77.3	217	2	STSH
9	197	75.8	215	2	STSH
10	197	75.8	217	1	STSH
11	197	75.8	217	2	STSH
12	161.5	62.1	216	1	STSH
13	160.5	61.7	190	2	STSH
14	159.5	61.3	190	1	STSH
15	159.5	61.3	190	2	STSH
16	159.5	61.3	190	2	STSH
17	159.5	61.3	216	1	STSH
18	159.5	61.3	216	1	STSH
19	159.5	61.3	216	2	STSH
20	159.5	61.3	216	2	STSH
21	159.5	61.3	216	2	STSH
22	159.5	61.3	216	2	STSH
23	159.5	61.3	216	2	STSH
24	159.5	61.3	216	2	STSH
25	150	57.7	190	1	STSH
26	148	56.9	191	2	STSH
27	146	56.2	163	2	STSH
28	144	55.4	190	2	STSH
29	144	55.4	216	2	STSH

30 142.5 54.8 217 1 STBO
31 142.5 54.8 217 1 STSH
32 142.5 54.8 217 1 STSH
33 142.5 54.8 217 2 S32682
34 140 53.8 216 2 S04929
35 132 50.8 190 2 A56916
36 132 50.8 215 2 A51188
37 126 49.2 195 2 A51250
38 126 49.2 215 2 JS0037
39 122 46.9 199 2 B32435
40 116 44.5 183 2 A60623
41 98.5 37.3 87 4 A67761
42 97 37.3 260 2 A51114
43 87 33.5 210 2 S69263
44 87 33.5 210 2 S69262
45 87 33.5 210 2 S02764

ALIGNMENTS

RESULT :

STBO

1: Somatotropin 2 precursor [validated] - human
2: Alternate names: growth hormone 1; hGH-N; pituitary somatotropin
3: Contains: growth hormone 5K peptide; somatotropin 1, long form; somatotropin 1, s
4: Species: Homo sapiens (man)
5: Date: 24-Apr-1984 #sequence_revision 10-Feb-1995 #text_change 08-Dec-2000
6: Accession: A93731; A32435; A93694; A94247; A90051; A93397; A93778; A91764; A90217
7: RefSeq: P.M.; Moore, D.D.; Goodman, H.M.
8: Nucleic Acids Res. 9, 3719-3730, 1981
9: A>Title: Human growth hormone DNA sequence and mRNA structure: possible alternative
10: A:Reference number: A93731; MUID:82014939; PMID:6269091
11: A:Accession: A93731
12: A:Molecule type: DNA
13: A:Residues: 1-217 <DEN>
14: A:Cross-references: GR:V00520
15: A:Note: the 20K short form somatotropin lacks residues 58-72 (32-46 in the active h
16: R:Chen, E.Y.; Liao, Y.C.; Smith, D.H.; Barrera-Saldana, H.A.; Gellinas, R.E.; Seebur
17: Genomics 4, 479-487, 1999
18: A>Title: The human growth hormone locus: nucleotide sequence, biology, and evolution
19: A:Reference number: A32435; MUID:89307277; PMID:2744760
20: A:Accession: A32435
21: A:Molecule type: DNA
22: A:Residues: 1-217 <CHE>
23: A:Cross-references: GB:J03071; NID:G183148; FIDN:AAA52549.1; PID:G183149
24: R:Roskam, W.; Rougeon, F.
25: Nucleic Acids Res. 7, 305-320, 1979
26: A>Title: Molecular cloning and nucleotide sequence of the human growth hormone stru
27: A:Reference number: A93694; MUID:80034477; PMID:386281
28: A:Accession: A93694
29: A:Molecule type: mRNA
30: A:Residues: 1-217 <ROS>
31: A:Cross-references: GR:V00519
32: A:Note: 35-pro was also found
33: R:Marshall, J.A.; Halliwell, R.A.; Baxter, J.D.; Goodman, H.M.
34: Science 205, 602-607, 1979
35: A>Title: Human growth hormone: complementary DNA cloning and expression in bacteria
36: A:Reference number: A94247; MUID:79203293; PMID:377496
37: A:Accession: A94247
38: A:Molecule type: mRNA
39: A:Residues: 1-217 <MAR>
40: R:Li, C.H.; Dixon, J.S.; Liu, W.K.
41: Arch. Biochem. Biophys. 133, 70-91, 1969
42: A>Title: Human pituitary growth hormone. XIX. The primary structure of the hormone.
43: A:Reference number: A90648; MUID:69289202; PMID:5810834
44: A:Contents: annotation
45: R:Li, C.H.; Dixon, J.S.
46: Arch. Biochem. Biophys. 146, 233-236, 1971
47: A>Title: Human pituitary growth hormone. XXII. The primary structure of the hormone
48: A:Reference number: A90051; MUID:72143935; PMID:5144027
49: A:Accession: A90051
50: A:Molecule type: protein

A:Residues: 27-94,96-217 <LIC>
 R:Niall, H.D.
 Nature New Biol. 250, 90-91, 1972.
 A:Title: Revised primary structure for human growth hormone.
 A:Reference number: A93397; NCID 7113765; PMID:5276046
 A:Accession: A93397
 A:Molecule type: protein
 A:Residues: 27-51 <NIA>
 R:Niall, H.D.; Honan, M.L.; Sayer, R.; Rosenblum, I.Y.; Greenwood, P.M.
 Proc. Natl. Acad. Sci. U.S.A. 68, 866-869, 1971
 A:Title: Sequences of pituitary and placental lactogenic and growth hormones: evolution.
 A:Reference number: A93778; NCID 71154968; PMID:5276524
 A:Accession: A93778
 A:Molecule type: protein
 A:Residues: 119-120;157-159 <N12>
 R:Niall, H.D.
 in Prolactin and Carcinogenesis, Proc. Fourth Tenev's Workshop: Prolactin, Grifflins, K.
 A:Title: The chemistry of the human lactogenic hormones.
 A:Reference number: A94427
 A:Contents: annotation; somatotropin revision
 R:Bewley, T.A.; Dixon, J.S.; Li, C.H.
 Int. J. Pept. Protein Res. 4, 281-287, 1972
 A:Title: Sequence comparison of human pituitary growth hormone, human chorionic somatotropin, and placental lactogen.
 A:Reference number: A91764; NCID:73092628; PMID:4675454
 A:Accession: A91764
 A:Molecule type: protein
 A:Residues: 27-217 <BEW>
 R:Lewis, G.J.; Bonewald, L.F.; Lewis, L.J.
 Biochem. Biophys. Res. Commun. 92, 511-516, 1980
 A:Title: The 20,000-dalton variant of human growth hormone: location of the amino acid
 A:Reference number: A93227; NCID:80130196; PMID:7336479
 A:Contents: somatotropin, 20K short variant
 A:Accession: A90217
 A:Molecule type: protein
 A:Residues: 46-57;73-80 <LEW>
 R:Chapman, G.E.; Rogers, K.M.; Brittain, T.; Bradshaw, R.A.; Pates, G.M.; Turner, D.; Ca
 J. Biol. Chem. 256, 2395-2403, 1981
 A:Title: The 20,000 molecular weight variant of human growth hormone: Preparation and se
 A:Reference number: A92311; NCID:6117361; PMID:7462247
 A:Contents: somatotropin, 20K short variant
 A:Accession: A92311
 A:Molecule type: protein
 A:Residues: 27-57;73-79 <CHA>
 R:Singh, R.N.P.; Saavey, B.K.; Lewis, L.J.; Lewis, G.J.
 J. Protein Chem. 2, 425-435, 1983
 A:Title: Human growth hormone peptide 1-45: Isolation from pituitary glands
 A:Reference number: A61456
 A:Accession: A61456
 A:Molecule type: protein
 A:Residues: 27-69 <SIN>
 R:Rebston, V.M.J.; Rao, I.B.; NG, F.
 Biol. Chem. Hoppe-Seyler 371, 423-431, 1990
 A:Title: Identification of the aspartamide structure in a previously-reported peptide
 A:Reference number: S09685; NCID:90334745; PMID:2138673
 A:Accession: S09685
 A:Molecule type: protein
 A:Residues: 27-34, L, 36-47 <KCB>
 R:De Vos, A.M.; Ullrich, M.; Kossiakoff, A.A.
 Science 255, 306-312, 1992
 A:Title: Human growth hormone and extracellular domain of its receptor: crystal structure
 A:Reference number: A41728; NCID:52296577; PMID:1545776
 A:Contents: annotation; X-ray crystallography, 2.8 angstroms
 A:Note: the structure of the complex with growth hormone receptor is described
 R:Gray, G.L.; Balridge, J.S.; McKown, K.S.; Heyneker, H.L.; Chang, C.N.
 Gene 39, 247-254, 1985
 A:Title: Periplasmic production of correctly processed human growth hormone in Escherich
 A:Reference number: I41126; NCID:8637393; PMID:391262
 A:Accession: I41549
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-26 <RES>
 A:Cross-references: GB:M14398; NCID:8183158; PIDN:AA52554.1; PID:g18159

Comment: The gene for this hormone is transcribed only in somatotrophic cells of 1
 Comment: About 90% of somatotropin is the 22K long form.
 C:Genetics:
 A:Gene: GH3:GH1
 A:Cross-references: GDB:111982; OMIM:136250
 A:Map position: 17q23.1-17q23.3
 A:Introns: 4/1; 57/3; 97/3; 152/3
 C:Superfamily: prolactin
 C:Keywords: alternative splicing; hormone; pituitary
 F:1-26/Domain: signal sequence; status predicted <SIG>
 F:27-217/Product: somatotropin 1, long form; status experimental <SOL>
 F:27-69/Product: growth hormone 5K peptide; status experimental <SKP>
 F:27-57,73-217/Product: somatotropin 1, short form; status experimental <SOS>
 F:79-191,208-215/Disulfide bonds: *Status experimental

Query Match 98.1%; Score 255; DB 1; Length 217;
 Best Local Similarity 100.0%; Pred. No. 2e-24; Indels 0; Gaps 0;
 Matches 48; Conservative 0; Mismatches 0;
 QY 2 FPTPLSLRLFDNAMLRAHRLHQLAFDTYQEFPEAYIPKEQKYSFLQNP 49
 DB 27 FPTPLSLRLFDNAMLRAHRLHQLAFDTYQEFPEAYIPKEQKYSFLQNP 74
 RESULT 2
 167410
 somatotropin - rhesus macaque
 K:Alternate names: growth hormone
 C:Species: Macaca mulatta (rhesus macaque)
 C:Date: 31-May-1996; #sequence_revision 31-May-1996; #text_change 16-Jul-1999
 C:Accession: I67410; A05094
 R:Golcs, T.G.; Durning, M.; Fisher, J.M.; Fowler, P.D.
 Endocrinology 133, 1744-1752, 1993
 A:Title: Cloning of four growth hormone/chorionic somatomammotropin-related compleme
 A:Reference number: I53267; NCID:94008724; PMID:8404617
 A:Accession: I67410
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-217 <RES>
 A:Cross-references: GB:I16556; NCID:9293114; PIDN:AA418842.1; PID:g293115
 R:Li, C.H.; Chung, D.; Lahn, H.W.; Stein, S.
 Arch. Biochem. Biophys. 245, 287-291, 1986
 A:Title: The primary structure of monkey pituitary growth hormone.
 A:Reference number: A05094; NCID:86129460; PMID:3080959
 A:Accession: A05094
 A:Molecule type: protein
 A:Residues: 27-99; 101-178; 180-217 <LIC>
 A:Note: the monkey species is not identified in the reference
 R:Ruben, M.S.
 Science 125, 883-884, 1957
 A:Title: Preparation of growth hormone from pituitaries of man and monkey.
 A:Reference number: A44774
 A:Contents: annotation; identification of source organism
 C:Superfamily: prolactin

Query Match 98.1%; Score 255; DB 2; Length 217;
 Best Local Similarity 100.0%; Pred. No. 2e-24; Indels 0; Gaps 0;
 Matches 48; Conservative 0; Mismatches 0;
 QY 2 FPTPLSLRLFDNAMLRAHRLHQLAFDTYQEFPEAYIPKEQKYSFLQNP 49
 DB 27 FPTPLSLRLFDNAMLRAHRLHQLAFDTYQEFPEAYIPKEQKYSFLQNP 74
 RESULT 3
 167410
 somatotropin 2 precursor - human
 K:Alternate names: growth hormone 2; growth hormone variant; hGH-V; placental somato
 N:Contents: somatotropin 2, long splice form; somatotropin 2, short splice form
 C:Species: Homo sapiens (man)
 C:Date: 17-Dec-1982; #sequence_revision 10-Feb-1995; #text_change 21-Jul-2000
 C:Accession: D32435; B28072; A01511; I52104; A60711
 R:Chen, E.Y.; Liao, Y.C.; Smith, D.H.; Barrera-Saldana, H.A.; Gellinas, R.E.; Seeburg

Genomics 4, 479-497, 1999
 A:Title: The human growth hormone locus: nucleotide sequence, biology, and evolution.
 A:Reference number: A12435; MUID:89307277; PMID:2744769
 A:Accession: D12435
 A:Molecule type: DNA
 A:Residues: 1-217 <GG>
 R:Cross-references: GB:103071; NID:9183148; PIDN:AAA52552.1; PID:9183152
 R:Cooke, N.E.; Ray, J.; Emery, T.G.; Lieberhaber, S.A.
 J. Biol. Chem. 263, 9001-9006, 1988
 A:Title: Two distinct species of human growth hormone variant mRNA in the human placenta
 A:Reference number: A92725; MUID:88243769; PMID:3379567
 A:Accession: B28072
 A:Molecule type: mRNA
 A:Residues: 1-217 <CC>
 R:Seeburg, P.H.
 DNA 1, 239-245, 1982
 A:Title: The human growth hormone gene family: nucleated the sequences show common descent
 A:Reference number: A01511; MUID:83182010; PMID:716905
 A:Accession: A01511
 A:Molecule type: DNA
 A:Residues: 1-34, P, 36-217 <SH>
 R:Ignotz, A.; Scippo, M.; Franckenne, F.; Hennion, B.
 Arch. Int. Physiol. Biochim. 85, 63-67, 1989
 A:Title: Cloning and nucleotide sequence of placental hGH-V cDNA.
 A:Reference number: 152104; MUID:89024984; PMID:246005
 A:Accession: 152104
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-217 <GG>
 R:Cross-references: GB:M38451; NID:9183179; PIDN:AAA5599.1; PID:9183181
 R:Franckenne, F.; Scippo, M.; Van Beuven, C.; Ignotz, A.; Hennion, B.
 J. Clin. Endocrinol. Metab. 71, 15-18, 1990
 A:Title: Identification of placental human growth hormone as the growth hormone-V gene
 A:Reference number: A50711; MUID:9037013; PMID:2199278
 A:Accession: A50711
 A:Molecule type: protein
 A:Residues: 27-44:45-57 <FRA>
 A:Experimental source: tissue placenta
 A:Note: partial glycosylation was demonstrated by lectin binding
 C:Comment: This gene is expressed by the placenta.
 C:Genetics:
 A:Gene: GDB:GH2
 A:Cross-references: GDB:110963; OMIM:139240
 A:Map position: 17q22-17q24
 A:Introns: 4/1; 57/3; 97/3; 152/3
 C:Superfamily: prolactin
 F:1-26/Domain: signal sequence #status predicted <SIG>
 F:27-217/Product: somatotropin 2, long splice form #status predicted <SIG>
 F:27-57.73-217/Product: somatotropin 2, short splice form #status predicted <SIG>
 F:79-191,208-215/Disulfide bonds: #status predicted
 F:166/Binding site: carbohydrate (Asn) (covalent) #status predicted
 Query Match 87.7%; Score 228; DB 1; Length 217;
 Best Local Similarity 91.7%; Pred. No. 4.5e-21;
 Matches 44; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 FPTPLSRFDNAMLRAHRLHQLAFDTYQFEFEAYIPKEKYSFLQNP 49
 DB 27 FPTPLSRFDNAMLRAHRLHQLAFDTYQFEFEAYIPKEKYSFLQNP 74
 RESULT 4
 STHUV2
 somatotropin 2 precursor, splice form 2 - human
 N:Alternate names: growth hormone variant-2; placental somatotropin form 2
 C:Species: Homo sapiens (man)
 C>Date: 30-Sep-1989 #sequence_revision 10-Pol-1995 #text_change 02-Sep-1997
 C:Accession: A28072
 R:Cooke, N.E.; Ray, J.; Emery, T.G.; Lieberhaber, S.A.
 J. Biol. Chem. 263, 9001-9006, 1988
 A:Title: Two distinct species of human growth hormone variant mRNA in the human placenta
 A:Reference number: A92725; MUID:88243769; PMID:3379567

A:Accession: A28072
 A:Molecule type: mRNA
 A:Residues: 1-256 <CC>
 A:Note: an alternative splice junction for intron 4 is used
 C:Genetics:
 A:Gene: GDB:GH2
 A:Cross-references: GDB:110963; OMIM:139240
 A:Map position: 17q22-17q24
 A:Introns: 4/1; 57/3; 97/3; 152/3
 C:Superfamily: prolactin
 C:Keywords: alternative splicing; hormone; placenta
 F:1-26/Domain: signal sequence #status predicted <SIG>
 F:27-256/Product: somatotropin 2 splice form 2 #status predicted <MAF>
 Query Match 87.7%; Score 228; DB 1; Length 256;
 Best Local Similarity 91.7%; Pred. No. 5.5e-21;
 Matches 44; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 FPTPLSRFDNAMLRAHRLHQLAFDTYQFEFEAYIPKEKYSFLQNP 49
 DB 27 FPTPLSRFDNAMLRAHRLHQLAFDTYQFEFEAYIPKEKYSFLQNP 74
 RESULT 5
 167478
 chorionic somatomammotropin-2 - rhesus macaque (fragment)
 C:Species: Macaca mulatta (rhesus macaque)
 C>Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 16-Jul-1999
 C:Accession: 167408
 R:Cooke, T.G.; Durning, M.; Fisher, J.M.; Fowler, P.D.
 Endocrinology 133, 1744-1752, 1993
 A:Title: Cloning of four growth hormone/chorionic somatomammotropin-related complex
 A:Reference number: 153267; MUID:94008724; PMID:8404617
 A:Accession: 167486
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-212 <RES>
 A:Cross-references: GB:L16553; NID:9293110; PIDN:AAA18840.1; PID:9293111
 C:Superfamily: prolactin
 Query Match 81.9%; Score 213; DB 2; Length 212;
 Best Local Similarity 78.7%; Pred. No. 3.4e-19;
 Matches 37; Conservative 9; Mismatches 1; Indels 0; Gaps 0;
 QY 3 HTPPLSRFLDNNALRAHRLHQLAFDTYQFEFEAYIPKEKYSFLQNP 49
 DB 23 PSVPLSRFLDNNALRAHRLHQLAFDTYQFEFEAYIPKEKHSLSMENP 69
 RESULT 6
 163267
 chorionic somatomammotropin-1 - rhesus macaque
 C:Species: Macaca mulatta (rhesus macaque)
 C>Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 16-Jul-1999
 C:Accession: 153267
 R:Cooke, T.G.; Durning, M.; Fisher, J.M.; Fowler, P.D.
 Endocrinology 133, 1744-1752, 1993
 A:Title: Cloning of four growth hormone/chorionic somatomammotropin-related complex
 A:Reference number: 153267; MUID:94008724; PMID:8404617
 A:Accession: 153267
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-217 <RES>
 A:Cross-references: GB:L16552; NID:9293108; PIDN:AAA18839.1; PID:9293109
 C:Superfamily: prolactin
 Query Match 81.9%; Score 213; DB 2; Length 217;
 Best Local Similarity 78.7%; Pred. No. 3.5e-19;
 Matches 37; Conservative 9; Mismatches 1; Indels 0; Gaps 0;
 QY 3 HTPPLSRFDNAMLRAHRLHQLAFDTYQFEFEAYIPKEKYSFLQNP 49
 DB 23 PSVPLSRFLDNNALRAHRLHQLAFDTYQFEFEAYIPKEKHSLSMENP 74

A:Reference number: A61283; MUID:91244006; PMID:2037148
 A:Accession: A61283
 A:Molecule type: protein
 A:Residues: 37-46 <NIG>
 A:Note: chorionamniotropon apparently copurified with placental catechol-O-methyltransferase
 R:Sherwood, L.M.; Handwerker, S.; McLaurin, W.D.; Langer, M.
 Nature New Biol. 233, 59-61, 1972
 A:Title: Amino-acid sequence of human placental lactogen.
 A:Reference number: A93401; MUID:72016313; PMID:5285363
 A:Contents: annotation
 R:Sherwood, L.M.; Handwerker, S.; McLaurin, W.D.; Langer, M.
 Nature New Biol. 233, 64, 1972
 A:Reference number: A93405
 A:Contents: annotation
 R:Schneider, A.B.; Kowalski, K.; Russell, J.; Sherwood, L.M.
 J. Biol. Chem. 254, 3782-3787, 1979
 A:Title: Identification of the interchain disulfide bonds of dimeric human placental lactogen.
 A:Reference number: A92251; MUID:79173083; PMID:5338159
 A:Contents: annotation; dimeric disulfide bonds
 R:Selby, M.J.; Partridge, A.; Baxter, J.D.; Bell, G.F.; Eberhardt, N.L.
 J. Biol. Chem. 259, 13131-13138, 1984
 A:Title: Analysis of a major human chorionic somatomammotropin gene. Evidence for two functional alleles.
 A:Reference number: 155229; MUID:85030426; PMID:6201192
 A:Accession: 155229
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-217 <RES>
 A:Cross-references: GR:K03401; NID:9181120; PIDN:AAA52155.1; PID:q51051
 R:Seeburg, P.H.; Shine, J.; Harada, J.A.; Gillich, A.; Goodman, H.
 Trans. Assoc. Am. Physicians 90, 109-116, 1977
 A:Title: Nucleotide sequence of a human gene coding for a polypeptide hormone.
 A:Reference number: 159658; MUID:78-63787; PMID:611657
 A:Accession: 159658
 A:Status: translated from GR/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 160-217 <R2>
 A:Cross-references: GB:M26118; NID:9181124; PIDN:AAA5521.1; PID:q51052
 C:Genetics:
 A:Gene: GRN/GSH
 A:Cross-references: GRN:112084; OMIM:150200
 A:Map position: 17q22-17q24
 A:Introns: 4/1; 5/1; 9/1; 152/4
 C:Superfamily: prolactin
 C:Keywords: hormone; placenta
 F:1-26/Domain: signal sequence #status experimental <SIG>
 F:27-217/Product: chorionamniotropon A #status experimental <EXP>
 F:79-191/Disulfide bonds: #status experimental
 F:208-215/Disulfide bonds: (in monomeric form) #status experimental
 F:208/Disulfide bonds: interchain (to 215 in dimeric form) #status experimental
 F:215/Disulfide bonds: interchain (to 208 in dimeric form) #status experimental

Query Match 75.8%; Score 197; DB 1; Length 217;
 Best Local Similarity 80.0%; Pred. No. 3.5e-17;
 Matches 36; Conservative 5; Mismatches 4; Indels 0; Gaps 0

QY 4 TPLSLFDNMLRAHRLHQLAFDTYQFEFAYIPKEOKYSFLQN 48
 DB 29 TVPLSLFDNMLRAHRLHQLAFDTYQFEFAYIPKEOKYSFLHD 73

RESULT 11
 E32435
 chorionamniotropon B precursor - human
 N:Alternate names: chorionic somatomammotropin 2
 C:Species: Homo sapiens (man)
 C:Date: 29-Dec-1989 #sequence_revision 29-Dec-1989 text_change 16-Jul-1999
 C:Accession: E32435
 R:Chen, E.Y.; Liao, Y.C.; Smith, D.H.; Herrera-Saidana, H.A.; Gelinas, A.E.; Seeburg, P.
 Genomics 4, 479-497, 1989
 A:Title: The human growth hormone locus: nucleotide sequence, organization, and evolution.
 A:Reference number: A32435; MUID:89307277; PMID:2744760
 A:Accession: E32435
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-217 <CHE>
 A:Cross-references: GB:J03071; NID:9183148; PIDN:AAA52553.1; PID:q183153
 C:Genetics:
 A:Gene: GDF/CSH2
 A:Cross-references: GRN:1119813; OMIM:118820
 A:Map position: 17q22-17q24
 C:Superfamily: prolactin

Query Match 75.8%; Score 197; DB 2; Length 217;
 Best Local Similarity 80.0%; Pred. No. 3.5e-17;
 Matches 36; Conservative 5; Mismatches 4; Indels 0; Gaps 0

QY 4 TPLSLFDNMLRAHRLHQLAFDTYQFEFAYIPKEOKYSFLQN 48
 DB 29 TVPLSLFDNMLRAHRLHQLAFDTYQFEFAYIPKEOKYSFLHD 73

RESULT 12
 STMS
 somatotropon precursor - mouse
 N:Alternate names: growth hormone
 C:Species: Mus musculus (house mouse)
 C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 text_change 28-May-1999
 C:Accession: E23911
 R:Linnar, D.L.H.; Talamantes, F.
 J. Biol. Chem. 260, 9574-9579, 1985
 A:Title: Nucleotide sequence of mouse prolactin and growth hormone mRNAs and expression of the complementary DNAs.
 A:Reference number: A92548; MUID:85261358; PMID:2991252
 A:Accession: E23911
 A:Molecule type: mRNA
 A:Residues: 1-216 <LIN>
 A:Cross-references: GB:X02891; GB:K03232; NID:951067; PIDN:CAA26650.1; PID:q51068
 C:Superfamily: prolactin
 C:Keywords: anterior pituitary; growth factor; hormone
 F:1-26/Domain: signal sequence #status predicted <SIG>
 F:27-216/Product: somatotropon #status predicted <SIG>
 F:74-189,206-214/Disulfide bonds: #status predicted

Query Match 62.2%; Score 161.5; DB 1; Length 216;
 Best Local Similarity 68.1%; Pred. No. 9.4e-13;
 Matches 32; Conservative 5; Mismatches 8; Indels 1; Gaps 1

QY 2 FTTPILSLFDNMLRAHRLHQLAFDTYQFEFAYIPKEOKYSFLQN 48
 DB 27 FTAMPILSLFDNMLRAHRLHQLAFDTYQFEFAYIPKEOKYSFLQN 72

RESULT 13
 PN0140
 somatotropon - sei whale
 N:Alternate names: growth hormone
 C:Species: Balenoptera borealis (sei whale)
 C:Date: 07-May-1993 #sequence_revision 07-May-1993 text_change 07-May-1999
 C:Accession: PN0140
 R:Yudaev, N.A.; Pankov, Y.A.; Bulatov, A.A.; Osipova, T.A.
 Biochimica 47, 1059-1069, 1982
 A:Title: Amino acid sequence of sei whale somatotropon.
 A:Reference number: PN0140; MUID:83000569; PMID:7115813
 A:Accession: PN0140
 A:Molecule type: protein
 A:Residues: 1-190 <YUD>
 A:Note: article in Russian with English abstract
 C:Superfamily: prolactin
 C:Keywords: growth factor; hormone
 F:52-163,180-188/Disulfide bonds: #status predicted

Query Match 61.7%; Score 160.5; DB 2; Length 190;
 Best Local Similarity 68.1%; Pred. No. 1.1e-12;
 Matches 32; Conservative 6; Mismatches 8; Indels 1; Gaps 1

QY 2 FTTPILSLFDNMLRAHRLHQLAFDTYQFEFAYIPKEOKYSFLQN 48
 DB 27 FTAMPILSLFDNMLRAHRLHQLAFDTYQFEFAYIPKEOKYSFLQN 72

DB 1 FPMPLSLPANAVALRQHLRQLAADTYKREFEAYTPKQKYSFLQN 46

RESULT 14

STHO

Somatotropin - horse
N:Alternate names: growth hormone
C:Species: Equus caballus (domestic horse)
C>Date: 13-Jul-1981 #sequence_revision:13-Jul-1981 #text_change:24-Aug-1996
C:Accession: A91772; A91395; A91383; A90240; A15134
R:Zakkin, M.M.; Poskus, E.; Langston, A.A.; Ferrara, P.; Santome, J.A.; Dellacha, J.A.
Int. J. Pept. Protein Res. 8, 435-444, 1976
A:Title: Primary structure of equine growth hormone.
A:Reference number: A91772; MUID:77005410; PMID:965151
A:Accession: A91772
A:Molecule type: protein
A:Residues: 1-190 <ZAK>
R:Zakkin, M.M.; Poskus, E.; Dellacha, J.A.; Paladini, A.C.; Santome, J.A.
FEBS Lett. 34, 353-355, 1973
A:Title: The amino acid sequence of equine growth hormone.
A:Reference number: A91395; MUID:74020362; PMID:4747849
A:Accession: A91395
A:Molecule type: protein
A:Residues: 1-190 <ZAK>
R:Zakkin, M.M.; Poskus, E.; Dellacha, J.A.; Paladini, A.C.; Santome, J.A.
FEBS Lett. 25, 77-82, 1972
A:Title: Amino acid sequences around the cysteine residues in equine growth hormone.
A:Reference number: A91383
A:Accession: A91383
A:Molecule type: protein
A:Residues: 42-69;157-190 <ZAK>
R:Oliver, L.; Hartree, A.S.
Biochem. J. 109, 19-24, 1968
A:Title: Amino acid sequences around the cysteine residues in horse growth hormone.
A:Reference number: A90240; MUID:68383390; PMID:4976109
A:Accession: A90240
A:Molecule type: protein
A:Residues: 176-190 <GLI>
C:Superfamily: prolactin
C:Keywords: hormone; pituitary
P:52-163,180-168/disulfide bonds: *status experimental

Query Match 61.3% Score 159.5; DB 2; Length 190;
Best Local Similarity 68.1% Pred. No. 1.4e-12;
Matches 32; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY 2 PPTPLSLPDNMLRAHRLHQLAFDTYQSEFEAYTPKQKYSFLQN 45

DB 1 FPMPLSLPANAVALRQHLRQLAADTYKREFEAYTPKQKYSFLQN 46

RESULT 15

JK0219

Somatotropin - African elephant
N:Alternate names: growth hormone
C:Species: Loxodonta africana (African elephant)
C>Date: 03-Aug-1992 #sequence_revision:03-Aug-1992 #text_change:15-Nov-1996
C:Accession: JK0219
R:Hulmes, J.D.; Miedel, M.C.; Li, C.H.; Fan, Y.C.H.
Int. J. Pept. Protein Res. 33, 356-372, 1989
A:Title: Primary structure of elephant growth hormone.
A:Reference number: JK0219
A:Accession: JK0219
A:Molecule type: protein
A:Residues: 1-190 <HUL>
A:Experimental source: pituitary gland
C:Superfamily: prolactin
P:1-190/Product: somatotropin *status experimental <MA7>

Query Match 61.3% Score 159.5; DB 2; Length 190;
Best Local Similarity 68.1% Pred. No. 1.4e-12;
Matches 32; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY 2 PPTPLSLPDNMLRAHRLHQLAFDTYQSEFEAYTPKQKYSFLQN 48
DB 1 FPMPLSLPANAVALRQHLRQLAADTYKREFEAYTPKQKYSFLQN 46

Search completed: September 15, 2003, 12:04:19
Job time : 9.07885 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2003, 11:54:00 : Search time 4.56631 Seconds
(without alignments)
504,635 Million cell updates/sec

Title: US-09-423-100-1

Perfect score: 260

Sequence: MFPT:PLSRKFDNM:KARY.....QRPFEAYIPKQKYAFQNP:49

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match (%)

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	255	98.1	217	1	SOMA_HUMAN
2	255	98.1	217	1	SOMA_MACRU
3	255	98.1	217	1	SOMA_PANTR
4	249	95.8	217	1	SOMA_CALFA
5	249	95.8	217	1	SOMA_SABR
6	236	90.8	217	1	SCM2_PANTR
7	228	87.7	217	1	SOMA_HUMAN
8	199	76.5	217	1	SOMA_MACRU
9	197	75.6	217	1	ELL_HUMAN
10	161.5	62.1	216	1	SOMA_MOUSE
11	160.5	61.7	190	1	SOMA_BALRO
12	159.5	61.3	190	1	SOMA_LOXAF
13	159.5	61.3	190	1	SOMA_VULVU
14	159.5	61.3	216	1	SOMA_CALFA
15	159.5	61.3	216	1	SOMA_FELCA
16	159.5	61.3	216	1	SOMA_HORSE
17	159.5	61.3	216	1	SOMA_MESAU
18	159.5	61.3	216	1	SOMA_PIG
19	159.5	61.3	216	1	SOMA_SABR
20	159.5	61.3	216	1	SOMA_RAT
21	159.5	61.3	217	1	SOMA_GAISE
22	159.5	61.3	217	1	SOMA_NYCPY
23	159.5	60.2	216	1	SOMA_MUSVI
24	155.5	59.8	190	1	SOMA_IAMPA
25	150	57.7	216	1	SOMA_YELCA
26	148	56.9	191	1	SOMA_CHEMF
27	145	55.8	215	1	SOMA_MONDO
28	145	55.8	215	1	SOMA_TRIVU
29	144	55.4	190	1	SOMA_ACIQU
30	144	55.4	190	1	SOMA_ACIQU
31	144	55.4	216	1	SOMA_CHICK
32	142.5	54.8	217	1	SOMA_BOVIN
33	142.5	54.8	217	1	SOMA_CERFEL

14	142.5	54.8	217	1	SOMA_SHEEP
35	142	54.6	217	1	SOMA_STRCA
36	140	53.8	190	1	SOMA_CRONO
37	140	53.8	216	1	SOMA_ANAPL
38	135.5	52.1	217	1	SOMA_BUBBU
39	132	50.8	215	1	SOMA_RANCA
40	125	48.1	211	1	SOMA_LEPOS
41	122	46.9	214	1	SOMA_XENLA
42	116	44.6	183	1	SOMA_PRIGL
43	112	43.1	206	1	SOMA_PROAN
44	111	42.7	213	1	SOMA_BUFMA
45	104	49.0	208	1	SOMB_XENLA

RESULT 1

SOMA_HUMAN

AC P01241: Q14405; Q16531; Q9HB22; Q90K57; Q90NL5; 217 AA.
21-JUL-1986 (Rel. 01, Created)
PT 1-MAR-1992 (Rel. 21, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Somatotropin precursor (Growth hormone) (GH) (GH-N) (Pituitary growth hormone) (Growth hormone :).
GN GH1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=80034477; PubMed=396281;
RA Koskam W., Hougden F.;
RT "Molecular cloning and nucleotide sequence of the human growth hormone structural gene".
RI Nucleic Acids Res. 7:305-320(1979).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=79203253; PubMed=377456;
RA Martini J.A., Hallett R.A., Baxter J.D., Goodman H.M.;
RT "Human growth hormone: complementary DNA cloning and expression in bacteria".
RI Science 205:622-627(1979).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND POSSIBLE ALTERNATIVE SPLICING.
RX MEDLINE=82014339; PubMed=6269091;
RA Denoto F.M., Moore D.D., Goodman H.M.;
RT "Human growth hormone DNA sequence and mRNA structure: possible alternative splicing".
RI Nucleic Acids Res. 9:3719-3730(1981).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=84182510; PubMed=7169009;
RA Seeburg P.H.;
RT "The human growth hormone gene family: nucleotide sequences show recent divergence and predict a new polypeptide hormone".
RI DNA 1:239-249(1982).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=89307277; PubMed=2744760;
RA Chen E.Y., Liao Y.C., Smith D.H., Barrera-Saldana H.A.;
RT "The human growth hormone locus: nucleotide sequence, biology, and evolution".
RI Genomics 4:479-497(1989).
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RX TISSUE=Pituitary;
RA Gu J., Huang O.-H., Li N., Xu S.-H., Han Z.-G., Fu G., Chen Z.;
RT "A novel gene expressed in human pituitary".
RI Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

RA SEQUENCE FROM N.A. (ISOFORM 4)
 RP TISSUE=Pituitary; PubMed=1334946;
 RX MEDLINE=26402571;
 RA Gu R.-M., Han Z.-G., Song H.-C., Peng Y.-D., Huang Q.-H., Ren G.-X.,
 RA Hu R.-M., Han Z.-G., Song H.-C., Peng Y.-D., Huang Q.-H., Ren G.-X.,
 RA Gu Y.-J., Huang Q.-H., Li Y.-P., Jiang G.-L., Fu G., Zhang Q.-B.,
 RA Gu B.-W., Dai M., Mao Y.-P., Guo G.-F., Kong F., Ye M., Zhang J.,
 RA Xu S.-H., Gu J., Shi J.-X., Jin W.-R., Zhao G.-K., Wu T.-Y.,
 RA Huang Q.-Y., Chen Z., Chen K.-D., Chen G.-L.,
 RT "Gene expression profiling in the human hypothalamus pituitary-adrenal
 RT axis and full-length cDNA cloning."
 RL Proc. Natl. Acad. Sci. U.S.A. 77:9543-9548(1980).
 RN [18]
 RP SEQUENCE OF 1-26 FROM N.A.
 RX MEDLINE=66137393; PubMed=39 2251;
 RA Gray G.L., Badierde J.S., McKown K.S., Heyneker H.L., Chang C.N.,
 RT "Periplasmic production of correctly processed human growth hormone in
 RT Escherichia coli: natural and bacterial signal sequences are
 RT interchangeable."
 RL Gene 39:247-254(1985).
 RN [19]
 RP SEQUENCE OF 27-217.
 RX MEDLINE=69289202; PubMed=58 0834;
 RA Li C.H., Dixon J.S., Liu W.K.,
 RT "Human pituitary growth hormone. XIX. The primary structure of the
 RT hormone."
 RL Arch. Biochem. Biophys. 133:70-91(1969).
 RN [10]
 RP SEQUENCE OF 27-217 AND REVISIONS.
 RX MEDLINE=72143935; PubMed=544027;
 RA Li C.H., Dixon J.S.,
 RT "Human pituitary growth hormone. 32. The primary structure of the
 RT hormone: revision."
 RL Arch. Biochem. Biophys. 146:233-246(1971).
 RN [11]
 RP REVISION.
 RX MEDLINE=71092028; PubMed=4675454;
 RA Bewley T.A., Dixon J.S., Li C.H.,
 RT "Sequence comparison of human pituitary growth hormone, human
 RT chorionic somatomammotropin, and ovine pituitary growth and
 RT lactogenic hormones."
 RL Int. J. Pept. Protein Res. 4:281-287(1972).
 RN [12]
 RP SEQUENCE OF 27-61 AND 103-134.
 RX MEDLINE=71139745; PubMed=5779045;
 RA Niall H.D.,
 RT "Revised primary structure of human growth hormone."
 RL Nature New Biol. 230:96-97(1971).
 RN [13]
 RP REVISIONS TO 119-120 AND 157-159.
 RX MEDLINE=71153968; PubMed=5279528;
 RA Niall H.D., Hodak M.E., Sauer R., Rosenblum I.Y., Greenwood F.C.,
 RT "Sequences of pituitary and placental lactogenic and growth hormones:
 RT evolution from a primordial peptide by gene reduplication."
 RL Proc. Natl. Acad. Sci. U.S.A. 68:866-869(1971).
 RN [14]
 RP REVISION.
 RX MEDLINE=71153968; PubMed=5279528;
 RA Niall H.D.,
 RT "The chemistry of the human lactogenic hormones."
 RL (In) Griffiths K. (eds.),
 RL Prolactin and carcinogenesis. Proc. Fourth Tancous Workshop Prolactin
 RL pp.13-20, Alpha Omega Alpha Press, Cardiff (1972).
 RN [15]
 RP SEQUENCE OF 27-79 (ISOFORM 2).
 RX MEDLINE=81117361; PubMed=7462247;
 RA Chapman G.E., Rogers K.M., Brittain J., Bradshaw E.A., Patay G.J.,
 RA Turner C., Cary P.D., Crane Robinson G.,
 RT "The 20,000 molecular weight variant of human growth hormone:
 RT Preparation and some physical and chemical properties."
 RL J. Biol. Chem. 256:2395-2401(1981).
 RN [16]
 RP SEQUENCE OF 46-80 (ISOFORM 2).
 RX MEDLINE=80130196; PubMed=7336479;
 Lewis U.G., Borewald L.F., Lewis L.J.,
 RT "The 20,000-dalton variant of human growth hormone: location of the
 RT amino acid deletions." Commun. 92:511-516(1980).
 RN [17]
 RP REAMINATION OF G.M.-163 AND ASN-179.
 RX MEDLINE=82052997; PubMed=7028740;
 RA Lewis U.G., Stapp R.N., Borewald L.F., Seavey B.K.,
 RT "Altered proteolytic cleavage of human growth hormone as a result of
 RT deamidation." J. Biol. Chem. 256:11645-11650(1981).
 RN [18]
 RP REVIEW.
 RX MEDLINE=99321812; PubMed=10393484;
 RA Bauwenn G.,
 RT "Growth hormone heterogeneity in human pituitary and plasma."
 RL Horm. Res. 51 Suppl. 1:2-6(1999).
 RN [19]
 RP 3D-STRUCTURE MODELING.
 RX MEDLINE=88190073; PubMed=3447173;
 RA Cohen F.R., Kuntz I.D.,
 RT "Prediction of the three-dimensional structure of human growth
 RT hormone."
 RL Proteins 2:162-166(1987).
 RN [20]
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
 RX MEDLINE=92196577; PubMed=1549776;
 RA de Vos A.M., Uitsch M., Kossiakoff A.A.,
 RT "Human growth hormone and extracellular domain of its receptor:
 RT crystal structure of the complex."
 RL Science 255:306-312(1992).
 RN [21]
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
 RX MEDLINE=95075462; PubMed=7984244;
 RA Somers W., Uitsch M., de Vos A.M., Kossiakoff A.A.,
 RT "The X-ray structure of a growth hormone-prolactin receptor complex."
 RL Nature 372:478-481(1994).
 RN [22]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RX Chantalat L., Chigadze N.Y., Jones N., Korber F., Navaza J.,
 RA Pavlovsk A.G., Wlodawer A.,
 RT "The crystal-structure of wild-type growth-hormone at 2.5-A
 RT resolution."
 RL Protein Pept. Lett. 2:333-340(1995).
 RN [23]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RX MEDLINE=97113023; PubMed=8943276;
 RA Sundstrom M., Lundqvist T., Roedin J., Giebel L.B., Milligan D.,
 RT "Crystal structure of an antagonist mutant of human growth hormone,
 RT G120R, in complex with its receptor at 2.9-A resolution."
 RL J. Biol. Chem. 271:32197-32203(1996).
 RN [24]
 RP VARIANT CYS-105.
 RX MEDLINE=99348093; PubMed=10391209;
 RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,
 RA Friedland L., Roife A., Warrington J., Lipshutz R., Daley G.C.,
 RA Lander E.S.,
 RT "Characterization of single-nucleotide polymorphisms in coding regions
 RT of human genes."
 RL Nat. Genet. 22:231-238(1999).
 RN [25]
 RP ERRATUM.
 RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,
 RA Friedland L., Roife A., Warrington J., Lipshutz R., Daley G.C.,
 RA Lander E.S.,
 RL Nat. Genet. 23:373-373(1999).
 CC -1- FUNCTION: Plays an important role in growth control. Its major
 CC role in stimulating body growth is to stimulate the liver and
 CC other tissues to secrete IGF-1. It stimulates both the
 CC differentiation and proliferation of myoblasts. It also stimulates

Query Match 98.1%, Score 255; DB 1; Length 217;
 Best Local Similarity 100.0%; Pred. No. 6.6e-25;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FFTPLSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEOKYSFLQNP 49
 |||||
 DB 27 FFTPLSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEOKYSFLQNP 74

RESULT 2

SOMA_MACMU STANDARD: PRT: 217 AA.
 AC P33693;
 DT 01-OCT-1994 (Rel. 27, Created)
 DI 01-OCT-1994 (Rel. 30, Last sequence update)
 DI 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Somatotropin precursor (Growth hormone) (GH) (GH-N) (Pituitary growth hormone) (Growth hormone 1).
 GN GH1.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Macaca.
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=94008724; PubMed=9404617;
 RA Colos T.G., Durning M., Fisher J.M., Fowler P.D.,
 RT "Cloning of four growth hormone/chorionic somatomotropin-related complementary deoxyribonucleic acids differentially expressed during pregnancy in the rhesus monkey placenta.";
 RT Endocrinology 133:1744-1752(1993).
 RN [2]
 RP SEQUENCE OF 27-217.
 RA MEDLINE=86129460; PubMed=3080959;
 RA Li C.H., Chung D., Lahn H.W., Stein S.;
 RI "The primary structure of monkey pituitary growth hormone.";
 RI Arch. Biochem. Biophys. 245:267-291(1986).
 CC -/- FUNCTION: Plays an important role in growth control. Its major role in stimulating body growth is to stimulate the liver and other tissues to secrete IGF-1. It stimulates both the differentiation and proliferation of myoblasts. It also stimulates amino acid uptake and protein synthesis in muscle and other tissues.

CC -/- SUBCELLULAR LOCATION: Secreted.
 CC -/- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).
 CC
 DR EMBL: L16556; AAA18842.1;
 DR PIR: I67410; I67410.
 DR HSSP: P01241; IAXI.
 DR InterPro: IPR001400; Somatotropin.
 DR Pfam: PF00103; hormone; 1.
 DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
 DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
 KW Hormone; Pituitary; Signal.
 FT SIGNAL 1 26
 FT CHAIN 27 217 SOMATOTROPIN.
 FT DISULFID 79 191 BY SIMILARITY.
 FT DISULFID 208 215 BY SIMILARITY.
 FT DISULFID 208 215 BY SIMILARITY.
 SQ SEQUENCE 217 AA; 24513 MW; 2C5180341EE24650 CRC64;

Query Match 98.1%; Score 255; DB 1; Length 217;
 Best Local Similarity 100.0%; Pred. No. 6.6e-25;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 FFTPLSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEOKYSFLQNP 49
 |||||
 DB 27 FFTPLSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEOKYSFLQNP 74

Query Match

98.1%; Score 255; DB 1; Length 217;

Best Local Similarity 100.0%; Pred. No. 6.6e-25;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FFTPLSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEOKYSFLQNP 49
 |||||
 DB 27 FFTPLSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEOKYSFLQNP 74

RESULT 3

SOMA_PANTR STANDARD: PRT: 217 AA.
 AC P58755;
 DI 28-FEB-2003 (Rel. 41, Created)
 DI 28-FEB-2003 (Rel. 41, Last sequence update)
 DI 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Somatotropin precursor (Growth hormone) (GH) (GH-N) (Pituitary growth hormone) (Growth hormone 1).
 GN GH1.
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
 OC NCBI_TaxID=9598;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Revol A., Esquivel D., Santiago D., Barreira-Saizda R.;
 RI "Independent duplication of the growth hormone gene in three Anthropoidea lineages.";
 RI Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
 CC -/- FUNCTION: Plays an important role in growth control. Its major role in stimulating body growth is to stimulate the liver and other tissues to secrete IGF-1. It stimulates both the differentiation and proliferation of myoblasts. It also stimulates amino acid uptake and protein synthesis in muscle and other tissues (by similarity).
 CC -/- SUBCELLULAR LOCATION: Secreted.
 CC -/- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AF374232; AAL72284.1;
 DR InterPro: IPR001400; Somatotropin.
 DR Pfam: PF00103; hormone; 1.
 DR PROSITE: PS00336; SOMATOTROPIN.
 DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
 DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
 KW Hormone; Pituitary; Signal.
 FT SIGNAL 1 26
 FT CHAIN 27 217 SOMATOTROPIN.
 FT DISULFID 79 191 BY SIMILARITY.
 FT DISULFID 208 215 BY SIMILARITY.
 FT DISULFID 208 215 BY SIMILARITY.
 SQ SEQUENCE 217 AA; 24843 MW; FEA295DE0518674 CRC64;

Query Match 98.1%; Score 255; DB 1; Length 217;
 Best Local Similarity 100.0%; Pred. No. 6.6e-25;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 FFTPLSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEOKYSFLQNP 49
 |||||
 DB 27 FFTPLSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEOKYSFLQNP 74

Best Local Similarity 100.0%; Pred. No. 6.6e-25;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FFTPLSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEOKYSFLQNP 49
 |||||
 DB 27 FFTPLSLRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEOKYSFLQNP 74

RESULT 4

SOMA_CALJA STANDARD: PRT: 217 AA.
 AC Q9GNB3;
 DI 28-FEB-2003 (Rel. 41, Created)
 DI 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Somatotropin precursor (Growth hormone).
 GN GH1.
 OS Callithrix jacchus (Common marmoset).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Platyrrhini; Callithricidae;
 OC Callithrix.
 OX NCBI_TaxID=9483;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wallis O.C., Wallis M.;
 RT "Cloning and characterization of a putative growth hormone encoding
 RL gene from the marmoset (Callithrix jacchus)."
 RL Submitted (AUG-2000) to the EMBL/GenBank/AAI databases.
 CC -!- FUNCTION: Plays an important role in growth control. Its major
 CC role is stimulating body growth. It stimulates both the
 CC differentiation and proliferation of myoblasts. It also stimulates
 CC other tissues to secrete IGF-1. It stimulates both the
 CC differentiation and proliferation of myoblasts. It also stimulates
 CC amino acid uptake and protein synthesis in muscle and other
 CC tissues (by similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AJ297563; CAC3481.1;
 CC HSP: P01241; I222.
 CC InterPro: IPR001400; Somatotropin.
 CC Pfam: PF00103; hormone; 1.
 CC PROSITE: PS00266; SOMATOTROPIN_1; 1.
 CC PROSITE: PS00338; SOMATOTROPIN_2; 1.
 CC Hormone; Pituitary; Signal.
 CC SIGNAL 1 26 BY SIMILARITY.
 CC CHAIN 27 217 SOMATOTROPIN.
 CC DISULFID 79 191 BY SIMILARITY.
 CC DISULFID 208 215 BY SIMILARITY.
 CC SEQUENCE 217 AA; 24959 MW; E:02151A12CE6:92 CRC64;
 CC
 CC Query Match 95.8%; Score 249; DB 1; Length 217;
 CC Best local Similarity 97.9%; Pred. No. 3.7e-24;
 CC Matches 47; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC 2 FPT:PLSLRFLNAMLRAHLRLHQLAFDTYQEFEEAYIPKEOKYSFLONP 49
 CC 27 FPT:PLSLRFLNAMLRAHLRLHQLAFDTYQEFEEAYIPKEOKYSFLONP 74
 CC
 CC RESULT 5
 CC SOMA_SAIIB STANDARD; PRT; 217 AA.
 CC AC P58343.
 CC DT 28-FEB-2003 (Rel. 41, Created)
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Somatotropin precursor (Growth hormone).
 GN GH1.
 OS Saimiri boliviensis boliviensis (Bolivian squirrel monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Saimiri.
 OX NCBI_TaxID=39432;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21265430; PubMed=11371582;
 RA Liu J.C., Makova K.D., Adkins R.M., Gibson S., Li W.H.;
 RT "Episodic evolution of growth hormone in primates and emergence of the
 RT species specificity of human growth hormone receptor."
 RL Mol. Biol. Evol. 18:945-953(2001).

CC -!- FUNCTION: Plays an important role in growth control. Its major
 CC role is stimulating body growth. It stimulates both the liver and
 CC other tissues to secrete IGF-1. It stimulates both the
 CC differentiation and proliferation of myoblasts. It also stimulates
 CC amino acid uptake and protein synthesis in muscle and other
 CC tissues (by similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF339560; AAK52287.1;
 CC InterPro: IPR001400; Somatotropin.
 CC Pfam: PF00103; hormone; 1.
 CC PROSITE: PS00336; SOMATOTROPIN.
 CC PROSITE: PS00266; SOMATOTROPIN_1; 1.
 CC PROSITE: PS00338; SOMATOTROPIN_2; 1.
 CC Hormone; Pituitary; Signal.
 CC SIGNAL 1 26 BY SIMILARITY.
 CC CHAIN 27 217 SOMATOTROPIN.
 CC DISULFID 79 191 BY SIMILARITY.
 CC DISULFID 208 215 BY SIMILARITY.
 CC SEQUENCE 217 AA; 24864 MW; 9515289992C529F7 CRC64;
 CC
 CC Query Match 95.8%; Score 249; DB 1; Length 217;
 CC Best local Similarity 97.9%; Pred. No. 3.7e-24;
 CC Matches 47; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC 2 FPT:PLSLRFLNAMLRAHLRLHQLAFDTYQEFEEAYIPKEOKYSFLONP 49
 CC 27 FPT:PLSLRFLNAMLRAHLRLHQLAFDTYQEFEEAYIPKEOKYSFLONP 74
 CC
 CC RESULT 6
 CC SOM2_PANIR STANDARD; PRT; 217 AA.
 CC AC P38757;
 CC DT 28-FEB-2003 (Rel. 41, Created)
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Growth hormone variant precursor (GH-V) (Placenta-specific growth
 CC hormone) (Growth hormone 2).
 GN GH2.
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
 OX NCBI_TaxID=9596;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Revol A., Esquivel D., Santiago D., Barrera-Saidana H.;
 RT "Independent duplication of the growth hormone gene in three
 RT Anthropoid lineages."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Plays an important role in growth control. Its major
 CC role is stimulating body growth. It stimulates both the
 CC other tissues to secrete IGF-1. It stimulates both the
 CC differentiation and proliferation of myoblasts. It also stimulates
 CC amino acid uptake and protein synthesis in muscle and other
 CC tissues.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Expressed in the placenta.
 CC -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to licensed@isb-sib.ch).

DR EMBL: AF374233; AAL72285.1.
 DR InterPro: IPR001400; Somatotropin.
 DR Pfam: PF00103; hormone; 1.
 DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
 DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
 KW Hormone; Placenta; Signal; Glycoprotein.
 FT SIGNAL 1 26 BY SIMILARITY.
 FT CHAIN 27 217 GROWTH HORMONE VARIANT.
 FT DISULFID 79 191 BY SIMILARITY.
 FT DISULFID 208 215 BY SIMILARITY.
 SQ SEQUENCE 217 AA: 24590 MW: 15924240756777E GR004;

Query Match 90.8% Score 236; D5 I; Length 217;
 Best Local Similarity 93.8% Prod No. 1.6e-22;
 Matches 45; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 2 FPTIPLSRLFNAMLRHLRLQLATYQYFEAYIFKECKYSFLQNP 49
 DB 27 FPTIPSLFQNMALRAFLRYQLAYTYQYFEAYIFKECKYSFLQNP 74
 RESULT 7
 ID SOM2_HUMAN STANDARD; PRT: 217 AA.
 AC P01242; P09587;
 DT 21-JUL-1986 (Rel. 01, Created)
 DI 28-FEB-2003 (Rel. 42, last sequence update)
 DI 15-SEP-2003 (Rel. 42, last annotation update)
 DE Growth hormone variant precursor (GH-V); (Placenta-specific growth
 DE hormone) (Growth hormone 2).
 GN GH2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1);
 RX MEDLINE=83182010; PubMed=7159509;
 RA Seeburg P.H.;

RT "The human growth hormone gene family: nucleotide sequences show
 RT recent divergence and predict a new polypeptide hormone.";
 RL DNA 1:239-249(1982).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RX MEDLINE=88243759; PubMed=379057;
 RA Cooke N.E., Ray J.J., Emery J.G., Liebhafner S.A.;

RT "Two distinct species of human growth hormone-variant mRNA in the
 RT human placenta predict the expression of novel growth hormone
 RT proteins".
 RL J. Biol. Chem. 263:9001-9005(1988).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=89024984; PubMed=2450050;
 RA Igout A., Scippo M.L., Franckenne P., Hennequin G.;

RT "Cloning and nucleotide sequence of placental hGH-V cDNA.";
 RL Arch. int. Physiol. Biochim. 96:63-67(1988).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89307277; PubMed=2744760;
 RA Chen E.Y., Liao Y.C., Smith D.H., Barrera-Saldana H.A.;

RA Gelinis R.E., Seeburg P.H.;

RT "The human growth hormone locus: nucleotide sequence, biology, and
 RT evolution.";
 RL Genomics 4:479-497(1989).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=22388257; PubMed=12177932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Berger J.G.,

KA Katscher R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Allschul S.F., Zeeberg B., Bietow K.E., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Blatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant I.L., Scheetz T.E.,
 RA Brownstein M.J., Jasin I.B., Yoshiyuki S., Carninci P., Pridge C.,
 RA Kana S.S., Loguc-Leco N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Rosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Wealey K.C., Hale S., Garcia A.M., Gay L.J., Huzyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Bickesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska J., Smalls D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [6]
 RP FEV-FW.
 RX MEDLINE=95321812; PubMed=10393484;
 RA Baumann G.;

RT "Growth hormone heterogeneity in human pituitary and plasma.";
 RL Horm. Res. 51 Suppl. 1:2-6(1999).
 RN [7]
 RP FUNCTION: Plays an important role in growth control. Its major
 RP role in stimulating body growth is to stimulate the liver and
 RP other tissues to secrete IGF-1. It stimulates both the
 RP differentiation and proliferation of myoblasts. It also stimulates
 RP amino acid uptake and protein synthesis in muscle and other
 RP tissues.

CC SUBUNIT: Monomer, dimer, trimer, tetramer and pentamer, disulfide-
 CC linked or non-covalently associated, in homopolymeric and
 CC heteropolymeric combinations. Can also form a complex either with
 CC GHBP or with the alpha2-macroglobulin complex.

CC SUBCELLULAR LOCATION: Secreted.

CC ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=2;

CC Name=1; Synonyms=GH-V1;

CC IsoId=P01242-1; Sequence=Displayed;

CC Name=2; Synonyms=GH-V2;

CC IsoId=P01242-2; Sequence=VSP_006203;

CC Note=No experimental confirmation available;

CC TISSUE SPECIFICITY: Expressed in the placenta.

CC SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to licensed@isb-sib.ch).

CC EMBL: K00470; AAB98619.1;
 CC EMBL: J03756; AAB59547.1;
 CC EMBL: J03756; AAB59548.1;
 CC EMBL: M3451; AAA35891.1;
 CC EMBL: J0307; AAA52352.1;
 CC EMBL: BC020760; AAH20760.1;
 CC PIR: A28072; SIHV2.
 CC PIR: D32435; SIHVJ.
 CC HSSP: P01241; 1A22.
 CC Gene: HGNC:4262; GH2.
 CC MIM: 139240;
 CC GO: GO:0005180; Peptide hormone; TAS.
 CC InterPro: IPR001400; Somatotropin.
 CC Pfam: PF00103; hormone; 1.
 CC PROSITE: PS00266; SOMATOTROPIN_1; 1.
 CC PROSITE: PS00338; SOMATOTROPIN_2; 1.
 KW Hormone; Placenta; Signal; Glycoprotein; Alternative splicing;
 KW Polymorphism. 1 26
 FT SIGNAL

EMBL: J00118; AAA36621.1; -
EMBL: BC002717; AA#02717.1; -
EMBL: BC005921; AAH05921.1; -
EMBL: BC020756; AAH20756.1; -
PIR: A26449; A26449.
PIR: C32435; LCHUC.
HSSP: POL24; IAZZ.
GeneID: HGNC:2440; CSHL.
GeneID: HGNC:2441; CSH2.
MIM: 150200; -
GO: GO:0007565; F.pregnancy; TAS.
InterPro: IPRO01403; Somatotropin.
Pfam: PF00183; hormone; 1.
PROSITE: PS00286; SOMATOTROPIN_1; 1.
PROSITE: PS00339; SOMATOTROPIN_2; 1.
Hormone: Placenta; Multigene family; Signal.
StoNAL 1 26
FT CHAIN 27 217 LACTOGEN.
FT DISULFID 75 191
FT DISULFID 208 215
FT DISULFID 208 208
FT DISULFID 208 215
FT DISULFID 208 215
FT VARIANT 3 3
FT VARIANT 104 105
FT CONFLICT 84 84
FT CONFLICT 95 95
FT CONFLICT 116 116
FT CONFLICT 134 134
FT CONFLICT 134 134
FT SEQUENCE 217 AA; 2359DC7A13F431 CRC64:
Query March 75.8%; Score 197; DB 1; Length 217;
Best Local Similarity 95.0%; Pred. No. 1,3e-17;
Matches 36; Conservativeness 5; Mismatches 4; Indels 0; Gaps

4 TPLSLRFEDNARLQALAFDTYQFEETAYTPKQYSFLON 48
||||| ||||| ||||| ||||| ||||| ||||| |||||
29 TVPLSRFEHANLQAHRHQALADTYQFEETIYPKQYSFLND 73

```

RESULT IN
SMA_MOUSE STANARD; PST; 216 AA.
FOE980:
01-JAN-1988 (Rel. 06, Created)
01-JAN-1988 (Rel. 06, Last sequence update)
28-FEB-2003 (Rel. 41, last annotation update)
Somatotropin precursor (growth hormone).
Gn CR GH.
Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_Taxid=10090;
SEQUENCE FROM N.A.
MEDLINE=85261358; PubMed=2991252;
Lizier D.I.H., Talamautes F.;
"Nucleotide sequence of mouse prolactin and growth hormone mRNAs and
expression of these mRNAs during pregnancy.";
J. Biol. Chem. 260:9574-9579(1985).
SEQUENCE FROM N.A.
MEDLINE=F2TDU; TISSUE=Liver;
MEELINP=96194803; PubMed=8647448;
Das P., Meyer L., Seyfert H.-M., Brockman G., Schwerin M.;
"Structure of the growth hormone-encoding gene and its promoter in
mice.";
Gene 159:209-213(1996).
```

-!- FUNCTION: Plays an important role in growth control. Its major role in stimulating body growth is to stimulate the liver and other tissues to secrete IGF-1. It stimulates both the differentiation and proliferation of myoblasts. It also stimulates

```

CC      amino acid uptake and protein synthesis in muscle and other
CC      tissues
CC      -!- SUBCELLULAR LOCATION: Secreted.
CC      -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC
CC      This Swiss-Prot entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL database.
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC      or send an email to license@sib-sib.ch).
CC
CC      EMBL: X02891; CAA26650.1;
CC      DR      EMBL: 245653; CAA86658.1;
CC      DR      PIR: B23911; STMS
CC      DR      HSSP: P01246; 1BS7.
CC      DR      MGD: MGI:95707; Gh.
CC      DR      InterPro: IPR001400; Somatotropin.
CC      DR      Pfam: PF00103; hormone; 1.
CC      DR      PROSITE: PS00266; SOMATOTROPIN_1; 1.
CC      DR      PROSITE: PS00338; SOMATOTROPIN_2; 1.
CC      KW      Hormone; Pituitary; Signal
CC      FT      SIGNAL 1 26 BY SIMILARITY.
CC      FT      CHAIN 27 216 SOMATOTROPIN.
CC      FT      DISULFID 78 189 BY SIMILARITY.
CC      FT      DISULFID 206 214 BY SIMILARITY.
CC      SQ      SEQUENCE: 216 AA; 24716 MW; 9866663AE25255FC CRC64;
CC
CC      Query Match 62.1%; Score 151.5; DB 1; Length 216;
CC      Best Local Similarity 68.1%; Pred. No. 3.6e-13;
CC      Matches 32; Conservative 6; Mismatches 8; Indels 1; Gaps 1;
CC
CC      QY      2 PPTPLSLRFDNALRAHLHQLAFTDYQEFEEAYIPKECKYSFLON 43
CC      DA      27 PPMPLSLSFNAVLRAQHILHQAADTYKEFERAYIPGQRYIS-IGN 72
CC
CC      RESULT 1;
CC      ID      SOMA_BALBO STANDARD; PRT: 190 AA.
CC      AC      P33092;
CC      DT      01-OCT-1993 (Rel. 27, Created)
CC      DI      01-OCT-1993 (Rel. 27, Last sequence update)
CC      D2      28-FEB-2003 (Rel. 41, Last annotation update)
CC      DE      Somatotropin (Growth hormone).
CC      GN      GHI.
CC      OS      Balaeonoptera borealis (Sei whale).
CC      OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC      CC      Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;
CC      CC      Balaeonopteridae; Balaeonoptera.
CC      OX      NCBI_TaxID=9768;
CC      RN      [1]
CC      RP      MEDLINE=83C00563; PubMed=7115813;
CC      RA      Yudaev N.A., Pankov Y.A., Gulatov A.A., Gaspova T.A.;
CC      RT      "Amino acid sequence of sei-whale somatotropin.";
CC      RL      Biochem Biophys Res Commun 199; 1059-1069 (1992)
CC      RN      [2]
CC      RP      PRELIMINARY PARTIAL SEQUENCE.
CC      RA      Gaspova T.A., Gulatov A.A., Pankov Y.A.;
CC      RT      "Structural studies of tryptic peptides from large cytosol protein
CC      RT      fragments of sei whale (Balaeonoptera borealis) somatotropin.";
CC      RL      Bioorg. Khim. 4; 1589-1599 (1978).
CC      CC      -!- FUNCTION: Plays an important role in growth control. Its major
CC      CC      role in stimulating body growth is to stimulate the liver and
CC      CC      other tissues to secrete IGF-1. It stimulates both the
CC      CC      differentiation and proliferation of myoblasts. It also stimulates
CC      CC      amino acid uptake and protein synthesis in muscle and other
CC      CC      tissues.
CC      -!- SUBCELLULAR LOCATION: Secreted.
CC      -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC      PIR: PNC140; PNC0140.

```

```

LR      HSSP: P01246; 1BS7.
LR      InterPro: IPR001400; Somatotropin.
LR      Pfam: PF00103; Hormone; 1.
LR      PROSITE: PS00266; SOMATOTROPIN_1; 1.
LR      PROSITE: PS00338; SOMATOTROPIN_2; 1.
LR      KW      Hormone; Pituitary.
LR      FT      DISULFID 52 163 BY SIMILARITY.
LR      FT      DISULFID 180 188 BY SIMILARITY.
LR      SQ      SEQUENCE: 190 AA; 21855 MW; 09BFF6D814A75D6 CRC64;
LR
LR      Query Match 61.7%; Score 150.5; DB 1; Length 190;
LR      Best Local Similarity 68.1%; Pred. No. 4.1e-13;
LR      Matches 32; Conservative 6; Mismatches 8; Indels 1; Gaps 1;
LR
LR      QY      2 PPTPLSLRFDNALRAHLHQLAFTDYQEFEEAYIPKECKYSFLON 48
LR      DA      1 PPMPLSLSFNAVLRAQHILHQAADTYKEFERAYIPGQRYIS-IGN 46
LR
LR      RESULT 12
LR      ID      SOMA_LOXAF STANDARD; PRT: 190 AA.
LR      AC      P20392;
LR      DT      01-FEB-1991 (Rel. 17, Created)
LR      DI      01-FEB-1991 (Rel. 17, Last sequence update)
LR      D2      24-FEB-2003 (Rel. 41, Last annotation update)
LR      DE      Somatotropin (Growth hormone).
LR      GN      GHI.
LR      OS      Loxodonta africana (African elephant).
LR      OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
LR      CC      Mammalia; Eutheria; Proboscidea; Elephantidae; Loxodonta.
LR      OX      NCBI_TaxID=9785;
LR      RN      [1]
LR      RP      SEQUENCE.
LR      RA      Hulmes J.D., Miedel M.C., Li C.H., Pan Y.C.E.;
LR      RT      "Primary structure of elephant growth hormone.";
LR      RL      Int. J. Pept. Protein Res. 33:368-372 (1989).
LR      CC      -!- FUNCTION: Plays an important role in growth control. Its major
LR      CC      role in stimulating body growth is to stimulate the liver and
LR      CC      other tissues to secrete IGF-1. It stimulates both the
LR      CC      differentiation and proliferation of myoblasts. It also stimulates
LR      CC      amino acid uptake and protein synthesis in muscle and other
LR      CC      tissues.
LR      -!- SUBCELLULAR LOCATION: Secreted.
LR      -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
LR      PIR: JK0219; JK0219.
LR      DR      HSSP: P01246; 1BS7.
LR      DR      InterPro: IPR001400; Somatotropin.
LR      DR      Pfam: PF00103; hormone; 1.
LR      DR      PROSITE: PS00266; SOMATOTROPIN_1; 1.
LR      DR      PROSITE: PS00338; SOMATOTROPIN_2; 1.
LR      KW      Hormone; Pituitary.
LR      FT      DISULFID 52 163 BY SIMILARITY.
LR      FT      DISULFID 180 188 BY SIMILARITY.
LR      SQ      SEQUENCE: 190 AA; 21761 MW; 058660813D8741F2 CRC64;
LR
LR      Query Match 61.3%; Score 150.5; DB 1; Length 190;
LR      Best Local Similarity 68.1%; Pred. No. 5.5e-13;
LR      Matches 32; Conservative 6; Mismatches 8; Indels 1; Gaps 1;
LR
LR      QY      2 PPTPLSLRFDNALRAHLHQLAFTDYQEFEEAYIPKECKYSFLON 48
LR      DA      1 PPMPLSLSFNAVLRAQHILHQAADTYKEFERAYIPGQRYIS-IGN 46
LR
LR      RESULT 13
LR      ID      SOMA_VULVU STANDARD; PRT: 190 AA.
LR      AC      P10766;
LR      DT      01-JUL-1989 (Rel. 11, Created)
LR      DI      01-JUL-1989 (Rel. 11, Last sequence update)
LR      D2      28-FEB-2003 (Rel. 41, Last annotation update)
LR      DE      Somatotropin (Growth hormone).

```

```

GN 1.
OS Vulpes vulpes (Red fox).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Vulpes.
OX NCBI_TaxID=9627;
RN [1]
RP SEQUENCE.
RC TISSUE=Pituitary;
RX MEDLINE=89254275; PubMed=2722401;
RA Li C.H., Izdebski J., Chung D.;
RI "Primary structure of fox pituitary growth hormone.";
RL Int. J. Pept. Protein Res. 33:70-72(1989).
CC -!- FUNCTION: Plays an important role in growth control. Its major
CC role in stimulating body growth is to stimulate the liver and
CC other tissues to secrete IGF-1. It stimulates both the
CC differentiation and proliferation of myoblasts. It also stimulates
CC amino acid uptake and protein synthesis in muscle and other
CC tissues.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC HSP: P01246; IBS.
CC InterPro: IPR01400; Somatotropin.
CC Pfam: PF00103; hormone; 1.
CC PROSITE: PS00266; SOMATOTROPIN_1; 1.
CC PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW Hormone; Pituitary.
FT DISULFID 52 163 BY SIMILARITY.
FT DISULFID 183 188 BY SIMILARITY.
SQ SEQUENCE 190 AA; 2174 MW; 14F37F9C1CB8692C CRC64;

Query Match 61.9%; Score 159.5; DB 1; Length 190;
Best Local Similarity 68.1%; Pred. No. 5.5e-13;
Matches 32; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

OY 2 FETPLSLRFLNAMLRAHLHQLADTYQFEFFAYIPKEQKYSFLQN 48
II III III III III III III III III III III III
Cb 1 FVAMPSSSLFANAVLRAQLHQLADTYKEFERAVIPEGQKYS-IQN 46

RESULT 14
SOMA_CANFA STANFORD; PRT; 216 AA;
AC P33711; O9QTQ;
DT 01-FEB-1994 (Rel. 25, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Somatotropin precursor (growth hormone).
GN GH1 OR GH.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94266166; PubMed=8626387;
RA Ascacio-Martinez J.A., Barrera-Saldana H.A.;
RI "A dog growth hormone cDNA codes for a mature protein identical to
RI pig growth hormone."
RL Gene 143:277-280(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA van Leeuwen I.S., Teske E., van Garderen E., Rutteman G.R., Mol J.A.;
RI "Extrapituitary growth hormone expression in the dog is initiated at
RI the normal pituitary transcription start site in the mammary gland and
RI at multiple upstream sites in lymphoid cells."
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=9337113; PubMed=10411306;
RA Lantinga-van Leeuwen I.S., Oudshoorn M., Mol J.A.;
RI "Canine mammary growth hormone gene transcription initiates at the
RI pituitary-specific start site in the absence of POU-1."

```

```

KL Mol. Cell. Endocrinol. 150:121-128(1999).
CC -!- FUNCTION: Plays an important role in growth control. Its major
CC role in stimulating body growth is to stimulate the liver and
CC other tissues to secrete IGF-1. It stimulates both the
CC differentiation and proliferation of myoblasts. It also stimulates
CC amino acid uptake and protein synthesis in muscle and other
CC tissues.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC This SW-SS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC ENBL: Z33057; CAAB0601.1;
CC ENBL: U92533; AAF21502.1;
CC ENBL: AF059071; AAD43366.1;
CC FIP: I46145; 146145.
CC HSP: P01246; IBS.
CC InterPro: IPR001400; Somatotropin.
CC Pfam: PF00103; hormone; 1.
CC PRINTS: PR00836; SOMATOTROPIN.
CC PROSITE: PS00266; SOMATOTROPIN_1; 1.
CC PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW Hormone; Pituitary; Signal.
FT SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 27 216 SOMATOTROPIN.
FT DISULFID 78 189 BY SIMILARITY.
FT DISULFID 206 214 BY SIMILARITY.
FT CONFLICT 4 4 S->G (IN REF. 1).
FT CONFLICT 7 7 N->T (IN REF. 1).
SQ SEQUENCE 216 AA; 24462 MW; ABAD:DD59FDAAED CRC64;

Query Match 61.9%; Score 159.5; DB 1; Length 216;
Best Local Similarity 68.1%; Pred. No. 6.3e-13;
Matches 32; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

OY 2 FETPLSLRFLNAMLRAHLHQLADTYQFEFFAYIPKEQKYSFLQN 48
II III III III III III III III III III III III
Cb 27 FVAMPSSSLFANAVLRAQLHQLADTYKEFERAVIPEGQKYS-IQN 72

RESULT 15
SOMA_FELCA STANFORD; PRT; 216 AA;
AC P46434;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Somatotropin precursor (growth hormone).
GN GH1.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary;
RX MEDLINE=96194906; PubMed=8654953;
RA Warren W.C., Bentley K.A., Bogosian G.;
RI "Cloning of the cDNAs coding for cat growth hormone and prolactin."
RL Gene 168:247-249(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary;
RX MEDLINE=9536973; PubMed=7642118;
RA Castro-Peralta F., Barrera-Saldana H.A.;
RI "Cloning and sequencing of cDNA encoding the cat growth hormone."
RL Gene 160:311-312(1995).

```

CC -1- FUNCTION: Plays an important role in growth control. Its major
CC role in stimulating body growth is to stimulate the liver and
CC other tissues to secrete IGF-1. It stimulates both the
CC differentiation and proliferation of myoblasts. It also stimulates
CC amino acid uptake and protein synthesis in muscle and other
CC tissues.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL, European
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (see <http://www.ist-sib.ch/omniprot/>
CC or send an email to license@sib-sb.ch).
CC -----
CC EMBL: G25973; AAA57294.1; .
CC EMBL: U13390; AAA96142.1; .
CC PIR: JC4632; JC4632.
CC HSP: P01246; IHSI.
CC InterPro: IPR001400; Somatotropin.
CC Pfam: PF06103; Hormone; 1.
CC PROSITE: PS00266; SOMATOTROPIN_1; 1.
CC PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW Hormone; Pituitary; Signal.
FT SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 27 216 SOMATOTROPIN.
FT DISULFID 78 189 BY SIMILARITY.
FT DISULFID 206 214 BY SIMILARITY.
FT CONFLICT 7 7 N -> T (IN REF. 2).
FT CONFLICT 26 26 T -> A (IN REF. 2).
FT CONFLICT 159 159 G -> A (IN REF. 2).
FT CONFLICT 181 181 L -> P (IN REF. 2).
SQ SEQUENCE 216 AA; 24454 MW; 65820239A7D292C6 CRC64;

Query Match 61.3%; Score 159.5; DB 1; length 216;
Best Local Similarity 68.1%; Pred. No. 6.3e-13;
Matches 32; Conservative 6; Mismatches 6; Indels 1; Gaps 1;

Cy 2 FPIIPLSRLFDNAMLRAHLHQLAFTYCFEEFAYIFKSKYFLQN 48
1 : ||| : |||| : || : ||||| : || : ||| :
Db 27 FPMPLSLFANAVIRACHLHQLAATYKEPERAYIDEGQVYS-IGN 72

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2003, 11:54:33 : Search time 19.1434 seconds
(without alignments)
660.520 Million cell updates/sec

Title: US-09-423-100-1

Perfect score: 260

Sequence: 1 MFPTPLSLRFDNMLRAHR.....QEFEEATYPRQKYSFLQNP 49

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 640525 seqs, 253652604 residues

Total number of hits satisfying chosen parameters: 810525

Minimum DB seq length: 6

Maximum DB seq length: 20000000

Post-processing: Minimum Match 10

Maximum Match 1000

Listing first 45 summaries

Database :

- 1: SP_ARCHAEA*
- 2: SP_BACTERIA*
- 3: SP_FUNGI*
- 4: SP_HUMAN*
- 5: SP_INVERTEBRATE*
- 6: SP_MAMMAL*
- 7: SP_MICE*
- 8: SP_ORNITHIFORMES*
- 9: SP_PLACENT*
- 10: SP_PLANT*
- 11: SP_PROTECT*
- 12: SP_VIRUS*
- 13: SP_VERTEBRATE*
- 14: SP_UNCLASSIFIED*
- 15: SP_VIRUS*
- 16: SP_BACTERIA*
- 17: SP_ARCHAEA*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	249	95.8	217	6	Q6WNEQ
2	228	87.7	245	4	Q14644
3	213	81.9	212	6	Q07368
4	213	81.9	217	6	Q07357
5	201	77.3	217	6	Q07369
6	197	75.8	217	4	Q14407
7	195	75.0	217	6	Q8WND9
8	170	65.4	217	6	Q8M174
9	160.5	61.7	216	11	Q70615
10	159.5	61.3	52	6	Q5TV91
11	159.5	61.3	216	6	Q8M173
12	159.5	61.3	216	6	Q8HYE5
13	155.5	59.8	216	11	Q9K2C3
14	154	59.2	216	11	Q5JRM4
15	152	58.5	178	6	Q65M75
16	149.5	57.5	204	6	Q95205

17	147	56.5	202	4	Q14643
18	146	56.2	178	6	Q95M76
19	144	55.4	190	11	Q9JKG0
20	143.5	55.2	192	6	Q9TU21
21	142.5	54.8	192	6	Q9TOW9
22	141	54.2	217	6	Q8M175
23	138.5	53.3	217	6	Q28957
24	138.5	53.3	217	6	Q9BEC0
25	138.5	53.3	217	6	Q9BEB9
26	137	52.7	40	6	Q9TR19
27	132	50.8	218	13	Q9PU72
28	128	49.2	195	13	Q91386
29	123.5	47.5	143	6	Q95240
30	122	46.9	149	4	Q14406
31	119	45.8	63	13	Q8QG85
32	107	41.2	53	6	Q19034
33	104	40.0	55	6	Q46474
34	101	38.8	167	4	P78451
35	101	38.8	200	13	Q8QF48
36	99.5	38.3	145	6	Q9BDR4
37	99	38.1	200	13	Q8QGJ1
38	99	38.1	200	13	Q8QFN2
39	96	36.9	50	13	Q9PRP5
40	95	36.5	200	13	Q8AV82
41	90	34.6	186	13	Q50283
42	87	33.5	168	13	Q8UVE2
43	87	33.5	168	13	Q9CW26
44	87	33.5	168	13	Q90W27
45	87	33.5	168	13	Q98SR7

ALIGNMENTS

RESULT 1

Q6WNEQ PRELIMINARY: PRT: 217 AA.
AC Q6WNEQ
BT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
ET 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Growth hormone.
GN GH-N.
CS Ateles geoffroyi (Black-headed spider monkey).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
NC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Ateles.
OX NCPI_TaxID:9509;
RN [1]
RP SEQUENCE FROM N.A.
SA Revol A., Esquivel L., Santiago D., Barrera-Saidana R.:
RT "Independent duplication of the growth hormone gene in three
RT Anthropoid lineages".
RL Submitted (APR-2001) to the EMBL/GenBank/DBSJ databases.
CR EMBL: AF374234; AAL7286.1; -
DR InterPro: IPR001400; Scatotropin.
DR Pfam: PF06103; Hormone; 1.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
SQ SEQUENCE 217 AA; 24894 MW; 425829FF41EEAAE6 CRC64;

Query Match 95.8%; Score 249; DB 6; Length 217;
Rest local similarity 97.9%; Pred. No. 1.2e-25;
Matches 47; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FPTPLSLRFDNMLRAHRLHOLAFDTYQEFEEATYPRQKYSFLQNP 49

QB 27 FPTPLSLRFDNMLRAHRLHOLAFDTYQEFEEATYPRQKYSFLQNP 74

RESULT 2

Q14644 PRELIMINARY: PRT: 245 AA.
AC Q14644

```

DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Placental growth hormone isoform B3H-V3 precursor.
GN HGH-V.
CS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homin.
OX NCBT_TaxID-9554;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Term placenta;
RX MEDLINE=98373737; PubMed 9709463;
RA Boustrenski C.L., Svensson B.A., Jansson T., Clark R.,
RA Carlsson L.M.S., Carlsson B.,
RI "Cloning of two novel growth hormone transcripts expressed in human
RI placenta.";
RL J. Clin. Endocrinol. Metab. 83:2878-2885(1998).
DR EMBL: AF006061; AAB71829.1; -.
DR HSSP: P01241; I422.
DR InterPro: IPR001400; Somatotropin;
DR Pfam: PF00103; Hormone;
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
KW SIGNAL.
FT SIGNAL.
SQ SEQUENCE 245 AA: 27101 NW: 14C076C075D5158 CRC64;
Query Match 87.7%; Score 228; DB 4; Length 245;
Best Local Similarity 91.7%; Pred. No. 9.5e-21;
Matches 44; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 FTTPLSRLFDNMLRAFLHQLAEDTYQEEAYIPKEKYSLQNP 49
DQ 27 FTTPLSRLFDNMLRAFLHQLAEDTYQEEAYIPKEKYSLQNP 74
DQ 27 FTTPLSRLFDNMLRAFLHQLAEDTYQEEAYIPKEKYSLQNP 74

RESULT 3
QY 007368 PRELIMINARY: PRT: 212 AA.
AC 007368;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Somatotropin 2 precursor (growth hormone 2) (Fragment).
CS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciidae; Macaca.
OX NCBT_TaxID-9544;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=94008724; PubMed-8404617;
RA Golos T.G., Durning M., Fisher J.M., Fowler P.D.,
RA "Cloning of four growth hormone/chorionic somatomotropin-related
RI complementary deoxyribonucleic acids differentially expressed during
RI pregnancy in the rhesus monkey placenta.";
RL Endocrinology 133:1744-1752(1995).
DR EMBL: L16553; AAA18840.1; -.
DR HSSP: P01241; IAXI.
DR InterPro: IPR001400; Somatotropin;
DR Pfam: PF00103; hormone;
DR PRINTS: PR00836; SOMATOTROPIN;
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
FT SIGNAL.
SQ SEQUENCE 212 AA: 24525 MW: 275C91104256E6F5 CRC64;
Query Match 87.9%; Score 213; DB 4; Length 212;
Best Local Similarity 78.7%; Pred. No. 8.7e-21;
Matches 37; Conservative 9; Mismatches 1; Indels 0; Gaps 0;

QY 3 FTTPLSRLFDNMLRAHRLHQLAEDTYQEEAYIPKEKYSLQNP 49
DQ 23 PSVPLSLFDHAMIQAHLHQLAEDTYQEEAYIPKEKHSMLNP 69
DQ 23 PSVPLSLFDHAMIQAHLHQLAEDTYQEEAYIPKEKHSMLNP 69

RESULT 4
QY 007369 PRELIMINARY: PRT: 217 AA.
AC 007369;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Somatotropin 2 precursor (growth hormone 2) (Fragment).
CS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciidae; Macaca.
OX NCBT_TaxID-9544;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Midpregnancy placenta;
RX MEDLINE=94008724; PubMed-8404617;
RA Golos T.G., Durning M., Fisher J.M., Fowler P.D.,
RA "Cloning of four growth hormone/chorionic somatomotropin-related
RI complementary deoxyribonucleic acids differentially expressed during
RI pregnancy in the rhesus monkey placenta.";
RL Endocrinology 133:1744-1752(1995).
DR EMBL: L16554; AAA18841.1; -.
DR HSSP: P01241; IAXI.
DR InterPro: IPR001400; Somatotropin;
DR Pfam: PF00103; hormone; 1.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
FT SIGNAL.
SQ SEQUENCE 217 AA: 24874 MW: F1EB6AF0BBAB1B185 CRC64;
Query Match 77.3%; Score 201; DB 6; Length 217;
Best Local Similarity 74.5%; Pred. No. 3.8e-19;

```



```

DE Growth hormone precursor.
OS Spalax leucodon ehrenbergi (Ehrenberg's mole rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Soturognathii; Muridae; Spalacidae;
OC Nannospalax.
OX NCBI_TaxID=30637;
RN [1];
RP SEQUENCE FROM N.A.
RX MEDLINE=99124645; PubMed=9724177;
RA Lioupias A., Nevo E., Wallis M.;
RT "Cloning and characterization of the gene encoding mole rat (Spalax
RT ehrenbergi) growth hormone."
RL C. Mol. Endocrinol. 22:25-36(1999).
DR EMBL: AJ005819; CAA06716.1; -.
DR HSSP: P01241; IAXI.
DR InterPro: IP001400; Somatotropin.
DR Pfam: PF00103; hormone; 1.
DR PRINTS: PR0836; SOMATOTROPIN.
DR PROSITE: PS02266; SOMATOTROPIN_1; 1.
DR PROSITE: PS02338; SOMATOTROPIN_2; 1.
KW Signal.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 256 GROWTH HORMONE.
FT CHAIN 27 256 GROWTH HORMONE.
SQ SEQUENCE 216 AA; 24627 MA; EEA88A523BA0ADFE CPC64;

Query Match 61.7%; Score 160.5; DB 1; Length 216;
Best Local Similarity 68.1%; Pred. No. 1.2e-13;
Matches 32; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY 2 FPTIPLSRFLDNAMLRHQLHQAADTYQEFEEAYIPKEOKYSFLON 48
DB 27 FPAMPLSSLFANAVLRACHLHQAADTYKEFERAYIPKEOKYS-ION 72

RESULT 10
Q9TV91 PRELIMINARY: PRT; 52 AA.
AC Q9TV91;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Growth hormone (Fragment).
GN Equus caballus (Horse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1];
RP SEQUENCE FROM N.A.
RX MEDLINE=99160443; PubMed=19951423;
RA Gaetano A.R., Pomp D., Murray S.D., Bowling A.L.;
RT "Comparative mapping of 16 equine type I genes assigned by somatic
RT cell hybrid analysis."
RL Mamm. Genome 10:271-276(1999).
DR EMBL: AF097589; AAC25992.1; -.
DR HSSP: P01241; IHGU.
DR InterPro: IP001400; Somatotropin.
DR Pfam: PF00103; hormone; 1.
FT NON_TER 1 52
FT NON_TER 52 52
SQ SEQUENCE 52 AA; 5835 MW; 20A9E9F5139F96AA CPC64;

Query Match 61.7%; Score 159.5; DB 6; Length 52;
Best Local Similarity 68.1%; Pred. No. 3.2e-14;
Matches 32; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

```

```

QY 2 FPTIPLSRFLDNAMLRHQLHQAADTYQEFEEAYIPKEOKYSFLON 48
DB 6 FPAMPLSSLFANAVLRACHLHQAADTYKEFERAYIPKEOKYS-ION 51

```

```

RESULT 11
Q8MI73

```

```

IU Q8MI73 PRELIMINARY: PRT; 216 AA.
AC Q8MI73;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Growth hormone precursor.
GN GH.
OS Delphinus delphis (Saddleback dolphin) (black sea dolphin).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti; Delphinidae;
OC Delphinus.
OX NCBI_TaxID=9722;
RN [1];
RP SEQUENCE FROM N.A.
RX TISSUE=Liver;
RA Manion D., Wallis O.G., Wallis M.;
RT "Cloning and characterisation of the GH gene from the common dolphin
RT (Delphinus delphis)."
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ492191; CAD37292.1; -.
DR InterPro: IP001400; Somatotropin.
DR Pfam: PF00103; hormone; 1.
DR PRINTS: PR0836; SOMATOTROPIN.
DR PROSITE: PS02266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW Signal.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 216 GROWTH HORMONE.
FT CHAIN 27 216 GROWTH HORMONE.
SQ SEQUENCE 216 AA; 24509 MW; 1EC467A84CCFEC2 CRC64;

Query Match 61.3%; Score 159.5; DB 6; Length 216;
Best Local Similarity 68.1%; Pred. No. 1.6e-13;
Matches 32; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY 2 FPTIPLSRFLDNAMLRHQLHQAADTYQEFEEAYIPKEOKYSFLON 48
DB 27 FPAMPLSSLFANAVLRACHLHQAADTYKEFERAYIPKEOKYS-ION 72

RESULT 12
Q8HYE5 PRELIMINARY: PRT; 216 AA.
AC Q8HYE5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Growth hormone precursor.
OS Alluropoda melanoleuca (Giant panda).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Ursidae; Alluropoda.
OX NCBI_TaxID=9646;
RN [1];
RP SEQUENCE FROM N.A.
RX TISSUE=Pituitary;
RA Liao M., Zhu M., Zhang A.;
RT "Cloning and expression of cDNA encoding growth hormone from
RT Alluropoda melanoleuca."
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF540936; AAN77228.1; -.
KW Signal.
FT SIGNAL 1 26 POTENTIAL.
FT SIGNAL 1 26 POTENTIAL.
SQ SEQUENCE 216 AA; 24393 MW; 44EC17EC44BCB056 CRC64;

Query Match 61.3%; Score 159.5; DB 6; Length 216;
Best Local Similarity 68.1%; Pred. No. 1.6e-13;
Matches 32; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY 2 FPTIPLSRFLDNAMLRHQLHQAADTYQEFEEAYIPKEOKYSFLON 48
DB 27 FPAMPLSSLFANAVLRACHLHQAADTYKEFERAYIPKEOKYS-ION 72

RESULT 13

```

```

Q9R2C3
ID Q9R2C3 PRELIMINARY: PRT: 216 AA.
AC Q9R2C3
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Growth hormone.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Scurionathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090.
RN [1]
RP SEQUENCE FROM N.A.
RA Nguyen T.N.K., Liehaber S.A.
RI "Mouse Growth Hormone Locus: Nucleotide Sequence and Phylogenetic
RI Analyses."
RL Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: U4362; AAC99988.1; .
DR HSP: P01241; IAXI.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone; 1.
DR PRINTS: PR00836; SOMATOTROPIN.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
SQ SEQUENCE 216 AA: 2482 MW: 42466.00 Da
Query Match 59.2%; Score 154; DB 11; Length 216;
Best Local Similarity 67.4%; Pred. No. 8.9e-13;
Matches 29; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 2 FTPTPLSLFDMALRAHRLHQLAFDTYCEFEAYIPKEOKYS 44
II III III III III III III III III III III III
DB 27 FPMPLSLFDMALRAHRLHQLAFDTYCEFEAYIPKEOKYS 44
II III III III III III III III III III III III

RESULT 14
Q9JMK4
ID Q9JMK4 PRELIMINARY: PRT: 216 AA.
AC Q9JMK4
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Growth hormone precursor.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hysticommatia; Cavidae; Cavi.
OX NCBI_TaxID=10141.
RN [1]
RP SEQUENCE FROM N.A.
RA Odorico L.M., Fuller P.J., Berington A.C.
RT "Cloning and sequence of guinea pig growth hormone (GH)."
RL Submitted (Feb-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF233653; AAF36409.1; .
DR HSP: P01241; IAXI.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone; 1.
DR PRINTS: PR00836; SOMATOTROPIN.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW Signal.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 216 GROWTH HORMONE.
SQ SEQUENCE 216 AA: 2482 MW: 45996.96 Da
Query Match 59.2%; Score 154; DB 11; Length 216;
Best Local Similarity 67.4%; Pred. No. 8.9e-13;
Matches 29; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 2 FTPTPLSLFDMALRAHRLHQLAFDTYCEFEAYIPKEOKYS 44
II III III III III III III III III III III III
DB 27 FPMPLSLFDMALRAHRLHQLAFDTYCEFEAYIPKEOKYS 44
II III III III III III III III III III III III

```

```

RESULT 15
Q9SM75
ID Q9SM75 PRELIMINARY: PRT: 178 AA.
AC Q9SM75
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Growth hormone (Fragment).
OS Tarsius bancanus (Western tarsier) (Horsfield's tarsier).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Tarsiidae; Tarsiidae; Tarsiidae.
OX NCBI_TaxID=94777.
RN [1]
RP SEQUENCE FROM N.A.
RA Liu J.C., Makova K.D., Adkins R.M., Gibson S., Li W.H.
RT "Episodic Evolution of Growth Hormone in Primates and Emergence of the
RT Species Specificity of Human Growth Hormone Receptor."
DR EMBL: AF339081; AAK62338.1; .
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone; 1.
DR PRINTS: PR00836; SOMATOTROPIN.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00266; SOMATOTROPIN_2; 1.
FT NON-TER 1 178
FT NON-TER 178 178
SQ SEQUENCE 178 AA: 20036 MW: 67208.75 Da
Query Match 58.5%; Score 152; DB 6; Length 178;
Best Local Similarity 65.1%; Pred. No. 1.3e-12;
Matches 28; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 2 FTPTPLSLFDMALRAHRLHQLAFDTYCEFEAYIPKEOKYS 44
II III III III III III III III III III III III
DB 8 FPMPLSLFDMALRAHRLHQLAFDTYCEFEAYIPKEOKYS 50
II III III III III III III III III III III III

Search completed: September 15, 2003, 12:03:27
Job time : 22.1434 secs

```

GenCore version 5.0.1
Copyright (c) 1993 - 2003 Computer Associates

OM protein - protein search, using sw model

Run on: September 15, 2003, 11:44:15 : Search time 14.6406 seconds
(without alignments)
225,942 Million cell updates/sec.

Title: US-09-423-100-2

Perfect score: 470

Sequence: 1 MEPTIPLSRLFDNMLRAHR.....NLEFLSLLLSWELRWQ 92

Scoring table: R20SUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 15 summaries

Database : A_Geneseq_10Jun03.*			
1:	/SIDSL/qcadata/geneseq/geneseq-emb1/AA1980.DAT.*		
2:	/SIDSL/qcadata/geneseq/geneseq-emb1/AA1981.DAT.*		
3:	/SIDSL/qcadata/geneseq/geneseq-emb1/AA1982.DAT.*		
4:	/SIDSL/qcadata/geneseq/geneseq-emb1/AA1983.DAT.*		
5:	/SIDSL/qcadata/geneseq/geneseq-emb1/AA1984.DAT.*		
6:	/SIDSL/qcadata/geneseq/geneseq-emb1/AA1985.DAT.*		
7:	/SIDSL/qcadata/geneseq/geneseq-emb1/AA1986.DAT.*		
8:	/SIDSL/qcadata/geneseq/geneseq-emb1/AA1987.DAT.*		
9:	/SIDSL/qcadata/geneseq/geneseq-emb1/AA1988.DAT.*		
10:	/SIDSL/qcadata/geneseq/geneseq-emb1/AA1989.DAT.*		
11:	/SIDSL/qcadata/geneseq/geneseq-emb1/AA1990.DAT.*		
12:	/SIDSL/qcadata/geneseq/geneseq-emb1/AA1991.DAT.*		
13:	/SIDSL/qcadata/geneseq/geneseq-emb1/AA1992.DAT.*		
14:	/SIDSL/qcadata/geneseq/geneseq-emb1/AA1993.DAT.*		
15:	/SIDSL/qcadata/geneseq/geneseq-emb1/AA1994.DAT.*		
16:	/SIDSL/qcadata/geneseq/geneseq-emb1/AA1995.DAT.*		
17:	/SIDSL/qcadata/geneseq/geneseq-emb1/AA1996.DAT.*		
18:	/SIDSL/qcadata/geneseq/geneseq-emb1/AA1997.DAT.*		
19:	/SIDSL/qcadata/geneseq/geneseq-emb1/AA1998.DAT.*		
20:	/SIDSL/qcadata/geneseq/geneseq-emb1/AA1999.DAT.*		
21:	/SIDSL/qcadata/geneseq/geneseq-emb1/AA2000.DAT.*		
22:	/SIDSL/qcadata/geneseq/geneseq-emb1/AA2001.DAT.*		
23:	/SIDSL/qcadata/geneseq/geneseq-emb1/AA2002.DAT.*		
24:	/SIDSL/qcadata/geneseq/geneseq-emb1/AA2003.DAT.*		

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	470	100.0	92	AAV42856	Human growth hormone
2	470	100.0	134	AAW92265	Human anti-angiogen
3	470	100.0	150	AAV42861	Chimeric protein
4	465	98.9	140	AAV91041	Human growth hormone
5	465	98.9	192	AAV90129	Human growth hormone
6	465	98.9	192	AAW92264	Human anti-angiogen
7	465	98.9	261	AAV92299	Human nerve growth
8	465	98.9	262	AAK11740	Human growth hormone
9	465	98.9	310	AAW03255	Fusion protein of

10	452	98.3	144	11	AAW05313	Segment of B-cell
11	452	98.3	262	7	AAV61033	Human beta-nerve g
12	450	97.9	138	9	AAV81226	Sequence of protei
13	450	97.9	191	18	AAO20110	Protein sequence o
14	450	97.9	191	20	AAV15809	Primary amino acid
15	450	97.9	191	20	AAV04396	Natural human 22kD
16	450	97.9	191	21	AAV04397	Mutant human 22kD
17	450	97.9	191	21	AAV78425	Human growth hormo
18	450	97.9	191	22	AAV17485	Human growth hormo
19	450	97.9	191	22	AAO17485	Human growth hormo
20	450	97.9	191	23	ABG94859	Human growth hormo
21	450	97.9	191	23	ABG94860	Human growth hormo
22	450	97.9	191	23	ABG94861	Human growth hormo
23	450	97.9	191	23	ABG94862	Human growth hormo
24	450	97.9	191	23	ABG94863	Human growth hormo
25	450	97.9	191	23	ABG94864	Human growth hormo
26	450	97.9	191	23	ABG94907	Human growth hormo
27	450	97.9	191	23	ABG94908	Human growth hormo
28	450	97.9	191	23	ABG94909	Human growth hormo
29	450	97.9	191	23	ABG94910	Human growth hormo
30	450	97.9	191	23	ABG94911	Human growth hormo
31	450	97.9	191	23	ABG94912	Human growth hormo
32	450	97.9	191	23	ABG94913	Human growth hormo
33	450	97.9	191	23	ABG94914	Human growth hormo
34	450	97.9	191	23	ABG94915	Human growth hormo
35	450	97.9	191	23	ABG94916	Human growth hormo
36	450	97.9	191	23	ABG94917	Human growth hormo
37	450	97.9	191	23	ABG94918	Human growth hormo
38	450	97.9	191	23	ABG94919	Human growth hormo
39	450	97.9	191	23	ABG94920	Human growth hormo
40	450	97.9	191	23	ABG94921	Human growth hormo
41	450	97.9	191	23	ABG94922	Human growth hormo
42	450	97.9	191	23	ABG94923	Human growth hormo
43	450	97.9	191	23	ABG94924	Human growth hormo
44	450	97.9	191	23	ABG94925	Human growth hormo
45	450	97.9	191	23	ABG94926	Human growth hormo

ALIGNMENTS

PEPIDE 1
AAV42856
ID : AAV42856 standard; protein: 92 AA.
AC AAV42856.
XX
XX 15-JAN-2000 (first entry)
XX Human growth hormone (hGH) N-terminal fragment #2.
XX Growth hormone; claptose; intramolecular; insulin; precursor;
XX Folding; conformation; chimeric protein; cleavable; recombinant;
XX Production; yield.
XX Homo sapiens.
XX
XX W05550302-A1.
XX
XX 07-OCT-1999.
XX
XX 31-MAR-1998; 98W0-CN00052.
XX
XX 31-MAR-1998; 98W0-CN00052.
XX
XX (TONG-) TONGHUA GANTECH BIOTECHNOLOGY LTD.
XX
XX Gen 2;
XX
XX WPI; 1559-610839/52.
XX
XX New chimeric proteins containing human growth hormone fragment, used
XX particularly for the production of human insulin -

XX PS Claim 5: Page 28: 46pp: English.

XX CC This sequence represents an N-terminal fragment of human growth hormone

CC CC (hGH) which is a component of a chimeric protein (AA42851) which also

CC CC contains a human insulin precursor (AA42850). The hGH portion of the

CC CC chimeric protein acts as an intramolecular chaperone (IMC) for the

CC CC insulin precursor, enabling it to fold correctly. A cleavable peptide

CC CC linker with a C-terminal Arg residue (AA42857) enables the hGH portion

CC CC of the chimeric protein to be removed after folding has taken place.

CC CC Production of recombinant human insulin via an 56H-proinsulin chimera

CC CC protein can provide human insulin with correctly linked cysteine bridges

CC CC with fewer necessary procedural steps, and hence resulting in a higher

CC CC yield of human insulin. The IMC sequences not only protect insulin

CC CC sequences from intracellular degradation by a microorganism host, but

CC CC also promote the folding of the fused insulin precursor, facilitate the

CC CC solubility of the fusion protein and decrease the intermolecular

CC CC interactions among the fusion proteins, thus allowing folding of the

CC CC fused insulin precursor at commercially useful high concentrations. The

CC CC procedural steps of cyanogen bromide cleavage, oxidative sulphydryl

CC CC and related purification steps can thus be eliminated, along with the

CC CC use of high concentrations of mercaptan or the use of hydrophobic

CC CC absorbent resins.

XX CC

XX CC Sequence 92 AA:

XX CC

XX CC Query Match 100.0%; Score 470; DB 20; Length 92;

XX CC Best Local Similarity 100.0%; Pred. No. 5.8e-40;

XX CC Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFTPIPLSRFDNAMLRAHRLHQLAFDTYQFEFAYTPKEQKYSFLONPOTSLSFSISIP 60

DB [|||||]||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1 MFTPIPLSRFDNAMLRAHRLHQLAFDTYQFEFAYTPKEQKYSFLONPOTSLSFSISIP 60

QY 61 IPSNREETOOKSNLELLRISLLLIQSWE:EPVO 92

DB [|||||]||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 61 TFSNREETOOKSNLELLRISLLLIQSWE:EPVO 92

RESULT 2

AAW92265

ID AAW92265 standard; Protein: 134 AA.

XX AC AAW92265:

XX DT 08-JUN-1999 (first entry)

XX DE Human anti-angiogenic peptide 15x hGH Met-1pro133.

XX KW Human; anti-angiogenic; prolactin; placental lactogen; hPL; angiogenesis;

XX KW growth hormone; hGH; hGH-V; capillary endothelial cell proliferation;

XX KW placental vascularisation; pregnancy; treatment; angiogenic disease;

XX KW tumour; inhibitor; malignant; angiofibroma; arteriovenous malformation;

XX KW arthritis; atherosclerotic plaques; corneal graft neovascularisation;

XX KW wound healing; proliferative retinopathy; macular degeneration; trachoma;

XX KW granulation; glaucoma; ocular; uveitis; fracture; Osler-Weber syndrome;

XX KW ulcer; leukaemia; reproductive disorder; contraceptive agent;

XX KW gene therapy; pre-eclampsia; intrauterine growth retardation;

XX KW placental dysfunction.

XX OS Homo sapiens.

XX PN W09851323-A1.

XX PD 19-NOV-1998.

XX PF 12-MAY-1998: 9AWO-USC96491.

XX PR 13-MAY-1997: 97US-0346394.

XX PA (REGC) UNIV CALIFORNIA.

XX CC Martini JA, Struman L, Taylor R, Weiner RI;

XX CC WEL: 1999-045192/04.

XX CC N-PSDB: AAXG1707.

XX CC New anti-angiogenic peptides - comprise N-terminal fragments of

XX CC human placental lactogen, human growth hormone, growth hormone

XX CC variant or human prolactin

XX CC Claim 4: Page 49-50: 57pp: English.

XX CC This invention describes novel human anti-angiogenic peptides derived

XX CC from 10 to 150 consecutive amino acids selected from the N-terminal end

XX CC of human placental lactogen (hPL), human growth hormone (hGH), growth

XX CC hormone variant (hGH-V), or human prolactin. Such peptides (i) inhibit

XX CC capillary endothelial cell proliferation and organisation (ii) inhibit

XX CC angiogenesis in chick chorioallantoic membrane and (iii) binds to at

XX CC least one specific receptor which does not bind an intact full length

XX CC hGH, hPL, prolactin or hGH-V. The invention also describes a method for

XX CC diagnosing a probable abnormality of placental vascularisation during

XX CC pregnancy. The peptides can be used for treating an angiogenic disease in

XX CC a subject, for inhibiting tumour formation or growth in a patient or for

XX CC modulating vascularisation of a patient's placenta. In particular, the

XX CC peptides can be used for preventing or treating e.g. malignant tumours,

XX CC angiofibroma, arteriovenous malformation, arthritic such as rheumatoid

XX CC arthritis, atherosclerotic plaques, corneal graft neovascularisation,

XX CC delayed wound healing, proliferative retinopathy such as diabetic

XX CC retinopathy, macular degeneration, granulations such as those occurring

XX CC in haemophilic joints, inappropriate vascularisation in wound healing

XX CC such as hypertrophic scars or keloid scars, neovascular glaucoma, ocular

XX CC tumour, uveitis, non-union fractures, Osler-Weber syndrome, psoriasis,

XX CC pyogenic glaucoma, retrolental fibroplasia, scleroderma, solid tumours,

XX CC Kaposi's sarcoma, trachoma, vascular adhesions, chronic varicose ulcers,

XX CC leukaemia, and reproductive disorders such as follicular and luteal cysts

XX CC and choriocarcinoma. They can also be used as contraceptive agents. DNA

XX CC encoding the peptides can be used in gene therapy. The measurement of

XX CC abnormal levels of N-terminal fragments of hGH, hGH-V, prolactin or hPL

XX CC can be used in assays for impairment of vascular development associated

XX CC with pre-eclampsia, intrauterine growth retardation, and placental

XX CC dysfunction.

XX CC

XX CC Sequence 134 AA:

XX CC

XX CC Query Match 100.0%; Score 470; DB 20; Length 134;

XX CC Best Local Similarity 100.0%; Pred. No. 8.8e-40;

XX CC Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFTPIPLSRFDNAMLRAHRLHQLAFDTYQFEFAYTPKEQKYSFLONPOTSLSFSISIP 60

DB [|||||]||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1 MFTPIPLSRFDNAMLRAHRLHQLAFDTYQFEFAYTPKEQKYSFLONPOTSLSFSISIP 60

QY 61 IPSNREETOOKSNLELLRISLLLIQSWE:EPVO 92

DB [|||||]||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 61 IPSNREETOOKSNLELLRISLLLIQSWE:EPVO 92

RESULT 3

AAW42851

ID AAW42851 standard; Protein: 150 AA.

XX AC AAW42851:

XX DT 19-JAN-2000 (first entry)

XX DE Chimeric protein, SH0 ID 7.

XX KW Insulin; precursor; growth hormone; chaperone; intramolecular;

XX KW folding; conformation; chimeric protein; cleavable; recombinant;

XX KW production; yield.

XX OS Synthetic.

XX OS Homo sapiens.

PN WO950302-A1.
 XX 07-OCT-1995.
 XX 31-MAR-1998; 98WO-CN00052.
 XX 31-MAR-1998; 98WO-CN00052.
 XX (TONG-) TONGHUA GANTECH BIOTECHNOLOGY LTD.
 XX Gan %;
 XX WPI; 1999-610339/52.
 XX New chimeric proteins containing human growth hormone fragment, used particularly for the production of human insulin.
 XX Claim 14; Page 30-31; 46pp; English.
 XX This sequence represents a chimeric protein, which contains an N-terminal fragment of human growth hormone (hGH) of the sequence given in AAY42856, a cleavable peptide linker (AAY42857), and a human insulin precursor comprising insulin A and B chains (AAY42859). The hGH portion of the chimeric protein acts as an intramolecular chaperone (IMC) for the insulin precursor, enabling it to fold correctly. The cleavable peptide linker has a C-terminal Arg residue which enables the hGH portion of the chimeric protein to be removed after folding has taken place. Production of recombinant human insulin via an hGH-proinsulin chimera protein can provide human insulin with correctly linked cysteine bridges with fewer necessary procedural steps, and hence resulting in a higher yield of human insulin. The IMC sequences not only protect insulin sequences from intracellular degradation by a microorganism host, but also promote the folding of the fused insulin precursor, facilitate the solubility of the fusion protein and decrease the intermolecular interactions among the fusion proteins, thus allowing foldings of the fused insulin precursor at commercially useful high concentrations. The procedural steps of cyanogen bromide cleavage, oxidative sulphatolysis and related purification steps can thus be eliminated, along with the use of high concentrations of detergent or the use of hydrophobic absorbent resins.
 XX Sequence 150 AA;
 XX Query Match 100.0%; Score 470; DB 20; Length 167;
 XX Best Local Similarity 100.0%; Pred. No. 10-39;
 XX Matches 92; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MPTTIPSLRFLFNAHLRAHRLHQLAFDTYQEEAYIPKEQKYSFLONPQTSLSFSESIP 60
 DB 1 MPTTIPSLRFLFNAHLRAHRLHQLAFDTYQEEAYIPKEQKYSFLONPQTSLSFSESIP 60
 QY 61 TPSNREETOQKSNLELLRSLILQSWLEPVQ 92
 DB 61 TPSNREETOQKSNLELLRSLILQSWLEPVQ 92
 RESULT 4
 AAP91041
 IC AAP91041 standard; protein, 140 AA.
 AC AAP91041;
 XX 14-DEC-1989 (first entry)
 XX Human growth hormone segment.
 XX Human growth hormone: fusion protein; thrombin; geriatric dementia; nervous disorders; human nerve factor.
 XX Homo sapiens (human).
 XX EP329175-A.
 XX

PJ 23-AUG-1989.
 XX 17-FEB-1989; 89EP-0102795.
 XX 15-FEB-1986; 86JP-0035042.
 XX (IOXI) TOSOH CORP.
 XX Ohtsuka E;
 XX WPI; 1989-243092/34.
 XX New human nerve growth factor gene encoding fusion protein
 XX - having cleavage site for thrombin, useful for treating geriatric dementia, etc.
 XX Disclosure; page 21; 38pp; English.
 XX Human growth hormone segment, used at the N-terminal of a fusion protein, which contains a thrombin recognition site, and human beta nerve growth factor (beta-NGF) at the C-terminal. Beta-NGF can be used to control geriatric dementia and other nervous disorders, and can be released from the fusion protein by incubation with thrombin (see AAN90577-8, AAP91034, AAP91299).
 XX Sequence 140 AA;
 XX Query Match 98.9%; Score 455; DB 10; Length 140;
 XX Best Local Similarity 98.9%; Pred. No. 3e-39;
 XX Matches 91; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MPTTIPSLRFLFNAHLRAHRLHQLAFDTYQEEAYIPKEQKYSFLONPQTSLSFSESIP 60
 DB 1 MPTTIPSLRFLFNAHLRAHRLHQLAFDTYQEEAYIPKEQKYSFLONPQTSLSFSESIP 60
 QY 61 TPSNREETOQKSNLELLRSLILQSWLEPVQ 92
 DB 61 TPSNREETOQKSNLELLRSLILQSWLEPVQ 92
 RESULT 5
 AAP90129
 IC AAP90129 standard; protein, 192 AA.
 AC AAP90129;
 XX 25-MAR-2003 (updated)
 XX 06-FEB-1996 (revised)
 XX 01-NOV-1989 (first entry)
 XX Human growth hormone.
 XX Human growth hormone: fusion protein; recombinant vector.
 XX Homo sapiens (human).
 XX JF01144981-A.
 XX 07-JUN-1989.
 XX 02-DEC-1987; 87JP-0304937.
 XX 02-DEC-1987; 87JP-0304937.
 XX (WAKT) WAKUNAGA SEIYAKU KK.
 XX WPI; 1989-209284/29.
 XX N-PSDB; AAN90269.
 XX Recombinant vector config. fusion protein - consisting of human growth hormone or deriv. ligated to foreign protein, for stability and high yield.
 XX

XX Disclosure: Fig 1: 19pp; Japanese.
 PS The invention consists of a vector contg. a fusion protein which is
 CC formed by ligating downstream of a promoter, hGH or a deriv. (prot.
 CC formed by substn. of Met-14 with leu) and a foreign protein.
 CC Stability of the vector in the host is greatly increased so the
 CC protein yield is higher.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 XX Sequence 192 AA:
 SQ
 Query Match 98.9%; Score 465; DB 20; Length 192;
 Best Local Similarity 98.9%; Pred. No. 4,20-39;
 Matches 91; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 XX
 QY 1 MFTPLSLRFDNAMLRAHRLHQLAFDTYQFEFAYIPKEQKYSLFQNPQTSLSFSESIP 60
 DB 1 MFTPLSLRFDNAMLRAHRLHQLAFDTYQFEFAYIPKEQKYSLFQNPQTSLSFSESIP 60
 QY 61 TPSNREETQOKSNLELLRISLLIQSWLEPVQ 92
 DB 61 TPSNREETQOKSNLELLRISLLIQSWLEPVQ 92
 RESULT 7
 AAFP91299
 ID AAFP91299 standard; protein: 261 AA.
 XX AAFP91299;
 XX
 XX 14-DEC-1989 (first entry).
 XX
 XX Human nerve growth factor and human growth hormone fusion protein.
 XX
 XX Human nerve growth factor; fusion protein; thrombin;
 KW geriatric dementia; nervous disorders; human growth hormone.
 XX
 XX Homo sapiens (human).
 XX
 XX Key Location/Qualifiers
 FT Region 1..140
 FT Region 141..143
 FT Region 144..261
 XX
 XX EF329175-A.
 XX
 XX 23-AUG-1989.
 XX
 XX 17-FEB-1989; 89EP-0102795.
 XX
 XX 19-FEB-1986; 89JP-0035042.
 XX
 XX (TQVJ) TOSOH CORP.
 XX
 XX Ohtsuka E;
 PI

XX Disclosure: Fig 1: 19pp; Japanese.
 PS The invention consists of a vector contg. a fusion protein which is
 CC formed by ligating downstream of a promoter, hGH or a deriv. (prot.
 CC formed by substn. of Met-14 with leu) and a foreign protein.
 CC Stability of the vector in the host is greatly increased so the
 CC protein yield is higher.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 XX Sequence 192 AA:
 SQ
 Query Match 98.9%; Score 465; DB 16; Length 192;
 Best Local Similarity 98.9%; Pred. No. 4,20-39;
 Matches 91; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 XX
 QY 1 MFTPLSLRFDNAMLRAHRLHQLAFDTYQFEFAYIPKEQKYSLFQNPQTSLSFSESIP 60
 DB 1 MFTPLSLRFDNAMLRAHRLHQLAFDTYQFEFAYIPKEQKYSLFQNPQTSLSFSESIP 60
 QY 61 TPSNREETQOKSNLELLRISLLIQSWLEPVQ 92
 DB 61 TPSNREETQOKSNLELLRISLLIQSWLEPVQ 92
 RESULT 6
 AAW92264
 ID AAW92264 standard; protein: 192 AA.
 XX AAW92264;
 XX
 XX 08-JUN-1999 (first entry)
 XX
 XX Human anti-angiogenic peptide hGH Met-1phel91.
 XX
 XX Human; anti-angiogenic; prolactin; placental lactogen; hPL; angiogenesis;
 KW growth hormone; hGH; hGH-V; capillary endothelial cell proliferation;
 KW placental vascularisation; pregnancy; treatment; angiogenic disease;
 KW tumour; inhibitor; malignant; angiofibroma; arteriovenous malformation;
 KW arthritis; atherosclerotic plaques; corneal graft neovascularisation;
 KW wound healing; proliferative retinopathy; macular degeneration; trachoma;
 KW granulation; glaucoma; ocular; uveitis; fracture; Osler-Weber syndrome;
 KW psoriasis; fibropia; scleroderma; Kaposi's sarcoma; vascular adhesion;
 KW ulcer; leukaemia; reproductive disorder; contraceptive agent;
 KW gene therapy; pre-eclampsia; intrauterine growth retardation;
 KW placental dysfunction.
 XX
 XX Homo sapiens.
 XX
 XX W09851323-A1.
 XX
 XX 19-NOV-1998.
 XX
 XX 12-MAY-1998; 98WO-US09491.
 XX
 XX 13-MAY-1997; 97US-0046394.
 XX
 XX (RECC) UNIV CALIFORNIA.
 XX
 XX Martial JA, Struman J, Taylor E, Weiner EJ;
 PI
 XX WPI: 1999-045192/04.
 XX
 XX N-PSDB: AAX01706.
 XX
 XX New anti-angiogenic peptides - comprise N-terminal fragments of
 PT human placental lactogen, human growth hormone, growth hormone
 PT variant or human prolactin.
 XX
 XX Example 3: Page 49; 97pp; English.
 XX
 XX This invention describes novel human anti-angiogenic peptides derived
 CC from 10 to 150 consecutive amino acids selected from the N-terminal end
 CC of human placental lactogen (hPL), human growth hormone (hGH), growth

AC AAR05313;
 XX
 DT 19-JUL-1990 (first entry)
 XX
 DE Segment of B-cell stimulatory factor-2 (IL-5).
 DE
 KW B-cell stimulatory factor-2; interleukin-5.
 XX
 OS Homo sapiens.
 XX
 PN JF02013375-A.
 XX
 PD 17-JAN-1990.
 XX
 PE 01-JUL-1988; 88JP-016255-5.
 XX
 PR 01-JUL-1988; 88JP-016255-5.
 XX
 PA (TOYO) TOYOKI CORP.
 XX
 XX WPI: 1990-062207/09.
 DR N-PSDR; AAR02528.
 XX
 XX Prepn. of human B-cell differentiation factor from specified DNA
 PT sequence segment, by recombinant DNA technique, gives protein of
 PT specified amino acid sequence.
 XX
 PS Disclosure; Page 5; 17pp; Japanese.
 XX
 XX The sequence encoding this protein can be fused with DNA encoding B cell
 CC differentiation factor (IL-5) and ligated into an expression vector for
 CC prodn. of a fusion protein.
 CC See also AAR0311.
 XX
 SQ Sequence 144 AA;

Query Match 98.3%; Score 452; DB 11; Length 144;
 Best Local Similarity 97.3%; Pred. No. 6.1e-39;
 Matches 90; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEFTIPLSLFDMNMLRAHRLHQLAFDYQVEFEAYIPKEQKYSFIQNFQTSLSSESIP 60
 DB 1 MEFTIPLSLFDMNMLRAHRLHQLAFDYQVEFEAYIPKEQKYSFIQNFQTSLSSESIP 60
 QY 61 TFSNRETCQKSNLELRISLLIQSWLFPVQ 92
 DB 61 TFSNRETCQKSNLELRISLLIQSWLFPVQ 92

RESULT 11
 AAP61033
 ID AAP61033 standard; Protein: 262 AA.
 XX
 AC AAP61033;
 XX
 DT 25-OCT-1991 (first entry)
 XX
 DE Human beta-nerve growth factor gene product.
 XX
 KW Beta-NGF; E.coli; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Protein 145..262

XX JF61205485-A.
 XX
 PD 11-SEP-1986.
 XX
 XX 09-MAR-1985; 85JP-0045773.
 XX
 PR 09-MAR-1985; 85JP-0045773.

XX
 PA (OTSUH/) OTSUKA E.
 XX
 LP WPI: 1986-281656/43.
 XX
 PT Gene segment of human nerve growth factor - used in prodn. of
 PT NGF-producing recombinant Escherichia strain.
 XX
 PS Claim 32; Page 482; 71pp; Japanese.
 XX
 CC The protein is a direct translation of the upstream tryptophan
 CC promoter-operator lacking its attenuation sequence and human
 CC beta-NGF sequence. The product may be efficiently expressed from a
 CC transformed E.coli expression system.
 CC See also AAN60816-7.
 XX
 SQ Sequence 262 AA;

Query Match 98.3%; Score 452; DB 7; Length 262;
 Best Local Similarity 97.3%; Pred. No. 1.2e-38;
 Matches 90; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEFTIPLSLFDMNMLRAHRLHQLAFDYQVEFEAYIPKEQKYSFIQNFQTSLSSESIP 60
 DB 1 MEFTIPLSLFDMNMLRAHRLHQLAFDYQVEFEAYIPKEQKYSFIQNFQTSLSSESIP 60
 QY 61 TFSNRETCQKSNLELRISLLIQSWLFPVQ 92
 DB 61 TFSNRETCQKSNLELRISLLIQSWLFPVQ 92

RESULT 12
 AAP61226
 ID AAP61226 standard; Protein: 138 AA.
 XX
 AC AAP61226;

XX
 DT 25-MAR-2003 (updated)
 DT 20-NOV-1990 (first entry)
 XX
 DE Sequence of protein with somatomedin-like activity.
 XX
 KW Growth hormone.
 XX
 OS Synthetic.
 XX
 PN JP63167798-A.
 XX
 PD 11-JUL-1986.

XX
 PF 29-DEC-1986; 86JP-0310177.
 XX
 PR 29-DEC-1986; 86JP-0310177.
 XX
 PA (TOYO) TOYO SODA MFG CO LTD.
 XX
 DR WPI: 1988-232632/33.
 DR N-PSDS; AAN81505.
 XX
 XX Polypeptide with somatomedin-like activity -
 PT by culturing bacterium transformed by plasmid contg. gene
 PT segment with specified DNA sequence
 XX
 XX Claim 2(1); Page 609; 9pp; Japanese.

XX
 XX The polypeptide (AAP61226) with somatomedin-like activity and the DNA
 CC (AAN81605) encoding it are claimed. A Met residual gp. may be added to
 CC the N-terminal. The polypeptide acts on the bone structure of mammals,
 CC including humans, to promote bone growth. The polypeptide has high
 CC production rate and is easily extracted from bacterial culture medium
 CC and refined for use as a bone growth accelerator.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX

```

SQ Sequence 135 AA.
Query Match 97.9%; Score 460; DB 15; Length 191;
Best Local Similarity 98.9%; Pred. No. 1.3e-38;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQFEFEAYIPKQKYSFLQNPQTSLSFSSEIPT 61
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQFEFEAYIPKQKYSFLQNPQTSLSFSSEIPT 60

QY 62 PSNREETOOKSNLELLRLISILLIQSWLEPVQ 92
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 PSNREETOOKSNLELLRLISILLIQSWLEPVQ 91

RESULT 13
ID AAO20110 standard; protein: 191 AA.
AC AAO20110;
XX
XX
XX 06-AUG-2002 (first entry)
DE Protein sequence of the hGH growth hormone cDNA.
KW Serum albumin-growth hormone fusion protein; growth hormones;
XX Down's syndrome.
XX Unidentified.
XX KR99076789-A.
XX 15-OCT-1999.
XX 25-JUN-1998; 96KR-070491A.
XX 30-DEC-1995; 95GB-002673A.
XX 19-DEC-1996; 96WO-GH0316A.
XX (DELTA ) DELTA BIOTECHNOLOGY LTD.
XX WPI: 1997-36369C/55.
XX N-PSDB: AAX99565.
XX
XX Serum albumin-growth hormone fusion protein - useful to treat growth
XX hormone related diseases, e.g. Down's syndrome
XX Disclosure: Fig 1; 2:pp; Karan.
XX
XX The invention relates to a serum albumin-growth hormone fusion protein
XX useful to treat growth hormone related diseases such as Down's syndrome.
XX This sequence represents a protein of the serum albumin-growth hormone
XX cDNA of the invention.
XX
XX Sequence 191 AA:
Query Match 97.9%; Score 460; DB 15; Length 191;
Best Local Similarity 98.9%; Pred. No. 1.3e-38;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQFEFEAYIPKQKYSFLQNPQTSLSFSSEIPT 61
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQFEFEAYIPKQKYSFLQNPQTSLSFSSEIPT 60

QY 62 PSNREETOOKSNLELLRLISILLIQSWLEPVQ 92
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 PSNREETOOKSNLELLRLISILLIQSWLEPVQ 91

RESULT 14
ID AAY15809 standard; protein: 191 AA.
XX
XX
XX 28-JUL-1999 (first entry)
DE Primary amine acid sequence of native human growth hormone.
XX
XX Detection; fluoresce; illegal misuse; growth substance; athlete;
XX domesticated farm animal; cattle; human growth hormone.
XX Homo sapiens.
XX WC9926069-A1.
XX
XX 27-MAY-1999.
XX
XX 15-NOV-1996; 98WO-GH03449.
XX
XX 14-NOV-1997; 97GB-G023955.
XX
XX (GENE-) GENERIC BIOLOGICALS LTD.
XX
XX Atkinson A, Murphy JP;
XX WPI: 1999-336072/28.
XX
XX Use of tagged exogenous polypeptide
XX Disclosure: Fig 1; 36pp; English.
XX
XX The specification describes a method of detecting an exogenously
XX administered substance from a naturally-occurring endogenous substance,
XX the exogenous substance being tagged so that it fluoresces differently
XX from the endogenous one at a suitable wavelength. The tagging may
XX consist of one or more substitutions in tagged growth hormone
XX selected from G40Y, F52Y, W86F, Y, L, I or V I103Y or I137Y;
XX The method is used to distinguish between exogenously administered
XX substances as compared to naturally-occurring endogenous substances.
XX Especially mentioned is the illegal misuse of growth substances by
XX athletes or in domesticated farm animals e.g. cattle. The present
XX sequence represents native human growth hormone which may be used
XX in the method of the invention.
XX
XX Sequence 191 AA:
Query Match 97.9%; Score 460; DB 20; Length 191;
Best Local Similarity 98.9%; Pred. No. 1.3e-38;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQFEFEAYIPKQKYSFLQNPQTSLSFSSEIPT 61
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQFEFEAYIPKQKYSFLQNPQTSLSFSSEIPT 60

QY 62 PSNREETOOKSNLELLRLISILLIQSWLEPVQ 92
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 PSNREETOOKSNLELLRLISILLIQSWLEPVQ 91

RESULT 15
ID AAY04396 standard; protein: 191 AA.
XX
XX AAY04396;
XX
XX 29-JUN-1999 (first entry)
DE Natural human 22kDa growth hormone.
XX
XX Human; 22kDa growth hormone; hGH; mutant; thrombin; resistance;
XX plasma; decomposition.
XX
XX Homo sapiens.
XX
XX JPI1092499-A.

```

```

XX 06-APR-1999.
PD
XX
XX 22-SEP-1997; 97JP-0275277.
PF
XX
XX 22-SEP-1997; 97JP-0275277.
PR
XX
XX (SUMO ) SUMITOMO SEIYAKU KK.
PA
XX
XX WPI; 1999-283567/24.
DR
XX
XX A human growth hormone mutant - with equivalent activity to natural
PI human growth hormone
XX
XX Example 1; Page 5-6; 10pp; Japanese.
PS
XX
XX The present invention describes a human growth hormone mutant in which
CC the 134th Arg and the 155th Thr are replaced respectively by Asp and Phe
CC in the 1st to the 191st amino acid sequence of natural type human 22 kDa
CC growth hormone (hGH) and which has a resistance against decomposition by
CC thrombin. The present sequence represents the natural hGH. Also
CC described are: (1) a hGH mutant in which the 134th Arg, the 155th Thr
CC and the 140th Lys are replaced respectively by Asp, Pro and Ala in the
CC amino acid sequence of natural type hGH and which has a resistance
CC against decomposition by thrombin and plasmin; and (2) a drug
CC preparation containing the above hGH mutant as the active component.
CC The mutant hGH shows an activity approximately equivalent to that of
CC natural type hGH and shows a high stability in blood and body fluid.
XX
XX Sequence 19; AA:
SQ
Query Match 97.9%; Score 460; DB 2; Length 191;
Best Local Similarity 99.9%; Pred. NO. 1,30-14;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 PFTPLSLRFLDANGLRAHRLHQGLADTYGEPDEAVIPKCKYSPFLGNQVLS-SFSSSPT 60
DB 1 PFTPLSLRFLDANGLRAHRLHQGLADTYGEPDEAVIPKCKYSPFLGNQVLS-SFSSSPT 60
QY 62 PSNRRTCKSKNLELRISLQLTQSWLEPVQ 92
DB 61 PSNRRTCKSKNLELRISLQLTQSWLEPVQ 91

```

Search completed: September 15, 2003, 12:00:57
Job time : 65.6308 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Computer Ltd.

OM protein : protein search, using sw model

Run on: September 15, 2003, 11:52:40 : Search time 21.109 seconds
(without alignments)
184,449 Million cell updates/sec

Title: US-09-423-100-2

Perfect score: 472

Sequence: 1 MFPIPLSRFDNAKRAHR.....NELLRISLLIQSLRWQ 52

Scoring table: BLDSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 423:0658 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 6

Maximum DB seq length: 2500000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA*

1: /cgn2_6/plodata/1/1aa/5A_COXH.pep.*
2: /cgn2_6/plodata/1/1aa/5A_COXH.pep.*
3: /cgn2_6/plodata/1/1aa/5A_COXH.pep.*
4: /cgn2_6/plodata/1/1aa/5A_COXH.pep.*
5: /cgn2_6/plodata/1/1aa/5A_COXH.pep.*
6: /cgn2_6/plodata/1/1aa/5A_COXH.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	465	98.9	192	US-08-293-382-1	Sequence 1, Appl
2	460	97.9	191	US-09-284-878-4	Sequence 5, Appl
3	460	97.9	194	US-08-383-621-4	Sequence 5, Appl
4	460	97.9	194	US-08-459-508-4	Sequence 5, Appl
5	460	97.9	217	US-08-589-029-10	Sequence 1, Appl
6	460	97.9	217	US-08-784-582-10	Sequence 1, Appl
7	460	97.9	217	US-08-785-271-10	Sequence 1, Appl
8	460	97.9	217	US-08-759-628-11	Sequence 1, Appl
9	460	97.9	217	US-09-284-878-1	Sequence 1, Appl
10	460	97.9	241	US-09-424-6205-25	Sequence 25, Appl
11	460	97.9	245	US-09-285-630-56	Sequence 66, Appl
12	460	97.9	274	US-08-784-582-71	Sequence 71, Appl
13	460	97.9	360	US-08-784-582-71	Sequence 71, Appl
14	454	96.6	191	US-09-465-461-1	Sequence 1, Appl
15	454	96.6	217	US-08-187-756C-4	Sequence 1, Appl
16	454	96.6	217	US-08-459-486-51	Sequence 51, Appl
17	454	96.6	217	US-08-459-658-51	Sequence 51, Appl
18	454	96.6	217	US-08-710-324A-4	Sequence 4, Appl
19	454	96.6	217	US-09-411-657-4	Sequence 4, Appl
20	453	96.4	400	US-09-420-819-37	Sequence 37, Appl
21	453	96.4	401	US-09-420-819-36	Sequence 36, Appl
22	447	95.1	191	US-08-800-215C-18	Sequence 18, Appl
23	445	94.7	191	US-08-800-215C-26	Sequence 16, Appl
24	445	94.7	191	US-08-800-215C-20	Sequence 20, Appl
25	364.5	77.6	176	US-08-791-728-1	Sequence 1, Appl
26	364.5	77.6	176	US-08-590-774-1	Sequence 1, Appl
27	358.5	76.3	176	US-08-791-728-2	Sequence 2, Appl

ALIGNMENTS

RESULT: 1

US-08-093-383-1

: Sequence 1, Application US/08093383

: Patent No. 5489529

: GENERAL INFORMATION:

: APPLICANT: DeBoer, Herman A.

: APPLICANT: Heyneker, Herbert L.

: APPLICANT: Seeburg, Peter H.

: TITLE OF INVENTION: DNA for Expression of Bovine Growth Hormone

: NUMBER OF SEQUENCES: 30

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: Genentech, Inc.

: STREET: 460 Point San Bruno Blvd

: CITY: South San Francisco

: STATE: California

: COUNTRY: USA

: ZIP: 94080

: COMPUTER READABLE FORM:

: MEDIUM TYPE: 5.25 inch, 360 kb floppy disk

: COMPUTER: IBM PC compatible

: OPERATING SYSTEM: PC-DOS/MS-DOS

: SOFTWARE: palin (Genentech)

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/08/093,383

: FILING DATE: 14-JUL-1993

: CLASSIFICATION: 435

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER: 07/419827

: FILING DATE: 28-NOV-1990

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER: 07/198824

: FILING DATE: 05-APR-1986

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER: 06/632361

: FILING DATE: 19-JUL-1984

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER: 06/303687

: FILING DATE: 16-SEP-1981

: ATTORNEY/AGENT INFORMATION:

: NAME: Johnston, Sean A.

: REGISTRATION NUMBER: PJ5,910

: REFERENCE/DOCKET NUMBER: 46C4

: TELECOMMUNICATION INFORMATION:

: TELEPHONE: 415/225-3562

: TELEFAX: 415/952-9881

: TELEX: 910/371-7168

: INFORMATION FOR SEQ ID NO: 1:

: SEQUENCE CHARACTERISTICS:

: LENGTH: 192 amino acids

: TYPE: amio acid

Sequence 2, Appl
Patent No. 5424199
Sequence 5, Appl
Sequence 5, Appl
Sequence 8, Appl
Sequence 4, Appl
Sequence 1, Appl
Sequence 2, Appl
Patent No. 5210180
Sequence 2, Appl
Sequence 2, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 3, Appl
Sequence 1, Appl
Sequence 2, Appl

28 358.5 76.4 176 4 US-08-990-774-2
29 340 72.3 168 6 5424199-3
30 323.5 71.2 198 1 US-08-187-756C-5
31 358.5 71.2 198 2 US-08-710-324A-5
32 343.5 71.2 198 4 US-09-411-657-5
33 306.5 65.2 191 1 US-08-468-824-8
34 304.5 64.9 191 1 US-07-963-331D-4
35 302.5 64.4 216 2 US-09-105-651-1
36 301.5 64.1 190 1 US-08-388-267C-2
37 301.5 64.1 190 4 US-09-277-720-2
38 301.5 64.1 191 6 5210180-1
39 301.5 64.1 193 1 US-07-621-197C-2
40 301.5 64.1 193 1 US-08-363-982-2
41 301.5 64.1 193 2 US-08-383-621-1
42 301.5 64.1 193 3 US-08-459-906-1
43 301.5 64.1 216 2 US-09-105-651-3
44 299.5 63.7 190 1 US-07-963-331D-3
45 299.5 63.7 191 1 US-07-922-523-1

```

; TOPLOGY: linear
US-08-093-383-1
;
; Query Match          98.9%   Score 465:   DP 1:   Length 192:
; Best Local Similarity 98.9%   Pred. No. 1.4e-51:
; Matches 91: Conservative 5: Mismatches 1: Indels 0: Gaps 0:
;
; QY 1 MFPTPLSRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFQNPQTSLSSESIP 60
;      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; DB 1 MFPTPLSRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFQNPQTSLSSESIP 60
;
; QY 61 TPSREETQOKSNLELRISLLIQSWLEPVQ 92
;      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; DB 61 TPSREETQOKSNLELRISLLIQSWLEPVQ 92
;
; RESULT 2
US-09-284-978-5
; Sequence 5, Application US/09284978
; Patent No. 6342175
; GENERAL INFORMATION:
; APPLICANT: Olazaran, Maria Guerrero
; APPLICANT: Salvada, Hugo Barrera
; APPLICANT: Salvado, Jose Maria Viader
; TITLE OF INVENTION: Genetically Modified Methylophilic P. pastoris Yeast for the
; FILE REFERENCE: Production and Secretion of the Human Growth Hormone
; CURRENT FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: US/09/284,878
; PRIOR FILING DATE: 1997-10-24
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 5
; LENGTH: 192
; TYPE: EST
; ORGANISM: Homo sapiens
US-09-284-978-5
;
; Query Match          97.9%   Score 460:   DB 4:   Length 191:
; Best Local Similarity 98.9%   Pred. No. 6e-51:
; Matches 90: Conservative 0: Mismatches 1: Indels 0: Gaps 0:
;
; QY 2 FPTPLSRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFQNPQTSLSSESIP 61
;      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; DB 1 FPTPLSRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFQNPQTSLSSESIP 60
;
; QY 62 PSNREETQOKSNLELRISLLIQSWLEPVQ 92
;      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; DB 61 PSNREETQOKSNLELRISLLIQSWLEPVQ 91
;
; RESULT 3
US-08-383-621-4
; Sequence 4, Application US/08383621
; Patent No. 5951972
; GENERAL INFORMATION:
; APPLICANT: Daley, Michael J.
; APPLICANT: Buckwalter, Brian L.
; APPLICANT: Cady, Susan M.
; APPLICANT: Shieh, Hong-Ming
; APPLICANT: Bohlen, Peter
; APPLICANT: Seddon, Andrew P.
; TITLE OF INVENTION: Stabilization of Somatotropins And Other
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. Estelle J. Tsevdos
; STREET: 1917 West Main Street, P.O. Box 60
; CITY: Stamford
; STATE: Connecticut
; COUNTRY: U.S.A.
; ZIP: 06904-0060

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/383,621
; FILING DATE: 06-FEB-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/766,142
; FILING DATE: 25-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsevdos, Estelle J.
; REGISTRATION NUMBER: 31,145
; REFERENCE/DOCKET NUMBER: 31,278-01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 203-321-2756
; TELEFAX: 203-321-2971
; TELEX: 203-710-474-4059
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 194 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-383-621-4
;
; Query Match          97.9%   Score 460:   DB 2:   Length 194:
; Best Local Similarity 98.9%   Pred. No. 6.1e-51:
; Matches 90: Conservative 0: Mismatches 1: Indels 0: Gaps 0:
;
; QY 2 FPTPLSRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFQNPQTSLSSESIP 61
;      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; DB 4 FPTPLSRFDNAMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFQNPQTSLSSESIP 63
;
; QY 62 PSNREETQOKSNLELRISLLIQSWLEPVQ 92
;      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; DB 64 PSNREETQOKSNLELRISLLIQSWLEPVQ 94
;
; RESULT 4
US-08-459-906-4
; Sequence 4, Application US/08459906
; Patent No. 6010995
; GENERAL INFORMATION:
; APPLICANT: Daley, Michael J.
; APPLICANT: Buckwalter, Brian L.
; APPLICANT: Cady, Susan M.
; APPLICANT: Shieh, Hong-Ming
; APPLICANT: Bohlen, Peter
; APPLICANT: Seddon, Andrew P.
; TITLE OF INVENTION: Stabilization of Somatotropins and Other
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Cyanamid Company
; STREET: Cte Cyanamid Plaza
; CITY: Wayne
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07470-8426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,906
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Darryl L.

```

```

? REGISTRATION NUMBER: 34,276
? REFERENCE/DOCKET NUMBER: 31,278-03
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 201-831-3247
? TELEFAX: 201-831-3305
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 194 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
US-08-453-906-4

Query Match          97.9%: Score 460; DB 3; Length 217;
Best Local Similarity 98.9%: Pred. No. 6,1e-51;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PPTPLSKLFNAMLRAHRLHQLAFTDYQEEYAIPEKQYSLQNPQTSLSESIPT 61
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4 PPTPLSKLFNAMLRAHRLHQLAFTDYQEEYAIPEKQYSLQNPQTSLSESIPT 63
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 62 PSNREETQOKSNLELLRISLLIQSWLEPVQ 92
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 64 PSNREETQOKSNLELLRISLLIQSWLEPVQ 94
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
US-08-589-028-10
? Sequence 10, Application US/C8589028
? Patent No. 6087129
? GENERAL INFORMATION:
? APPLICANT: Newgard, Christopher B.
? APPLICANT: Halban, Philippe A.
? APPLICANT: No. 6087129minston, Karl D.
? APPLICANT: Clark, Samuel A.
? APPLICANT: Thigpen, Aulice E.
? APPLICANT: Quade, Christian
? APPLICANT: Kruse, Fred
? TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
? NUMBER OF SEQUENCES: 50
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Arnold, White & Durkee
? STREET: P. O. Box 4433
? CITY: Houston
? STATE: TX
? COUNTRY: USA
? ZIP: 77210-4433
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: IBM PC compatible
? SOFTWARE: Patent In Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/589,028
? FILING DATE: Concurrently Herewith
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Highlander, Steven J.
? REGISTRATION NUMBER: 37,642
? REFERENCE/DOCKET NUMBER: US/08/589,028
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (512) 418-3000
? TELEFAX: (512) 474-7577
? INFORMATION FOR SEQ ID NO: 10:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 217 amino acids
? TYPE: amino acid
? STRANDEDNESS:
? TOPOLOGY: linear
US-08-589-028-10

Query Match          97.9%: Score 460; DB 3; Length 217;
Best Local Similarity 98.9%: Pred. No. 6,1e-51;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PPTPLSKLFNAMLRAHRLHQLAFTDYQEEYAIPEKQYSLQNPQTSLSESIPT 61
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 27 PPTPLSKLFNAMLRAHRLHQLAFTDYQEEYAIPEKQYSLQNPQTSLSESIPT 86
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

US-08-784-382-10
? Sequence 10, Application US/08784582
? Patent No. 6110707
? GENERAL INFORMATION:
? APPLICANT: Newgard, Christopher B.
? APPLICANT: Halban, Philippe A.
? APPLICANT: No. 6110707minston, Karl D.
? APPLICANT: Clark, Samuel A.
? APPLICANT: Thigpen, Aulice E.
? APPLICANT: Quade, Christian
? APPLICANT: Kruse, Fred
? TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
? NUMBER OF SEQUENCES: 79
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Arnold, White & Durkee
? STREET: P. O. Box 4433
? CITY: Houston
? STATE: Texas
? COUNTRY: USA
? ZIP: 77210
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/784,582
? FILING DATE: Concurrently Herewith
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? PRIOR APPLICATION NUMBER: US 60/028,427
? FILING DATE: 15-OCT-1996
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/589,028
? FILING DATE: 15-JAN-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Highlander, Steven J.
? REGISTRATION NUMBER: 37,642
? REFERENCE/DOCKET NUMBER: US/08/589,028
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 512/418-3000
? TELEFAX: 512/474-7577
? INFORMATION FOR SEQ ID NO: 10:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 217 amino acids
? TYPE: amino acid
? STRANDEDNESS:
? TOPOLOGY: linear
US-08-784-582-10

Query Match          97.9%: Score 460; DB 3; Length 217;
Best Local Similarity 98.9%: Pred. No. 7,1e-51;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PPTPLSKLFNAMLRAHRLHQLAFTDYQEEYAIPEKQYSLQNPQTSLSESIPT 61
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 27 PPTPLSKLFNAMLRAHRLHQLAFTDYQEEYAIPEKQYSLQNPQTSLSESIPT 86
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Qy 62 PSNREETQOKSNLELLRISLLLIQSWLEPVQ 92
 Db 87 PSNREETQOKSNLELLRISLLLIQSWLEPVQ 117

RESULT 7

US-08-785-271-10
 : Sequence 10, Application US/23785271
 : Patent No. 6194176
 : GENERAL INFORMATION:
 : APPLICANT: Newgard, Christopher H.
 : APPLICANT: Haibian, Philippe A.
 : APPLICANT: No. 5194176 Inventor: Karl D.
 : APPLICANT: Clark, Samuel A.
 : APPLICANT: Thigpen, Anice E.
 : APPLICANT: Quade, Christan
 : APPLICANT: Kruse, Fred
 : TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
 : TITLE OF INVENTION: SECRETORY CELL LINES
 : NUMBER OF SEQUENCES: 56
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Arnold, White & Barker
 : STREET: P.O. Box 4433
 : CITY: Houston
 : STATE: Texas
 : COUNTRY: USA
 : ZIP: 77210
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent in Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/785,271
 : FILING DATE: Concurrently Herewith
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/589,028
 : FILING DATE: 19-JAN-1996
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Highlander, Steven L.
 : REGISTRATION NUMBER: 37,542
 : REFERENCE/DOCKET NUMBER: 07SR:513
 : TELEPHONE: 512/418-3000
 : TELEFAX: 512/474-7577
 : INFORMATION FOR SEQ ID NO: 10:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 217 amino acids
 : TYPE: amino acid
 : STRANDEDNESS: linear
 : TOPOLOGY: linear
 : US-08-785-271-10

Query Match 97.9%; Score 460; DB 2; Length 217;
 Best Local Similarity 98.9%; Pred. No. 7,1e-51;
 Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FPTPLSRLEFDNMLRAHLRHLQAFDTYQERFEAYIPKEQKYSFLQNPQTSLSSESPT 61

Db 27 FPTPLSRLEFDNMLRAHLRHLQAFDTYQERFEAYIPKEQKYSFLQNPQTSLSSESPT 86

Qy 62 PSNREETQOKSNLELLRISLLLIQSWLEPVQ 92
 Db 87 PSNREETQOKSNLELLRISLLLIQSWLEPVQ 117

RESULT 8

US-08-759-628-11
 : Sequence 11, Application US/28759628
 : Patent No. 6225446
 : GENERAL INFORMATION:
 : APPLICANT: Altman, Scott W.

APPLICANT: Brock, Fernando L.
 APPLICANT: Bazan, J. Fernando
 APPLICANT: Kasteleiner, Robert A.
 TITLE OF INVENTION: MUTATIONAL VARIANTS OF MAMMALIAN PROTEINS
 NUMBER OF SEQUENCES: 11
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: DNAX Research Institute
 STREET: 901 California Avenue
 CITY: Palo Alto
 STATE: California
 COUNTRY: USA
 ZIP: 94304-1104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/759,628
 FILING DATE: 05-DEC-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/006,574
 FILING DATE: 06-DEC-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Ching, Edwin P.
 REGISTRATION NUMBER: 34,090
 REFERENCE/DOCKET NUMBER: DX05520
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-852-9398
 TELEFAX: 415-496-1200
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 217 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FEATURE:
 NAME/KEY: Peptide
 LOCATION: 32..53
 FEATURE:
 NAME/KEY: Peptide
 LOCATION: 94..115
 FEATURE:
 NAME/KEY: Peptide
 LOCATION: 133..153
 FEATURE:
 NAME/KEY: Peptide
 LOCATION: 192..210
 OTHER INFORMATION: /note= *The peptides above are
 : OTHER INFORMATION: depicted in Figure 1*
 : US-08-759-628-11

Query Match 97.9%; Score 460; DB 3; Length 217;
 Best Local Similarity 98.9%; Pred. No. 7,1e-51;
 Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FPTPLSRLEFDNMLRAHLRHLQAFDTYQERFEAYIPKEQKYSFLQNPQTSLSSESPT 61

Db 27 FPTPLSRLEFDNMLRAHLRHLQAFDTYQERFEAYIPKEQKYSFLQNPQTSLSSESPT 86

Qy 62 PSNREETQOKSNLELLRISLLLIQSWLEPVQ 92
 Db 87 PSNREETQOKSNLELLRISLLLIQSWLEPVQ 117

RESULT 9

US-09-284-878-1
 : Sequence 1, Application US/09284878
 : Patent No. 6342375
 : GENERAL INFORMATION:
 : APPLICANT: Olazaran, Martha Guerrero

```

? APPLICANT: Seidara, Hugo Barreira
? APPLICANT: Salvado, Jose Maria Viador
? TITLE OF INVENTION: Genetically Modified Methylotrophic P. pastoris Yeast for the
? TITLE OF INVENTION: Production and Secretion of the Human Growth Hormone
? FILE REFERENCE: 1829.0010000
? CURRENT APPLICATION NUMBER: US/09/244,878
? PRIOR APPLICATION NUMBER: 1999-07-21
? PRIOR FILING DATE: 1997-10-24
? NUMBER OF SEQ ID NOS: 3
? SOFTWARE: Patentin Ver. 2.1
? SEQ ID NO 1
? LENGTH: 217
? TYPE: PRT
? ORGANISM: Homo sapiens
US-09-284-878-1

Query Match 97.9%; Score 460; DB 4; Length 217;
Best Local Similarity 98.9%; Pred. No. 7.1e-51;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FTPIPLSLFDNMLRAHRLHQLAFDTYOEFEEAYIPKEQKYSFLONPQTSLSFSESPT 61
Db 27 FTPIPLSLFDNMLRAHRLHQLAFDTYOEFEEAYIPKEQKYSFLONPQTSLSFSESPT 61
QY 62 PSNREETQCKSNLELIRISLLLIQSWLEPVQ 92
Db 87 PSNREETQCKSNLELIRISLLLIQSWLEPVQ 117

RESULT 10
US-09-424-6203-25
? Sequence 25, Application US/994246203
? Patent No. 5391595
? GENERAL INFORMATION:
? APPLICANT: HANIL SYNTHETIC FIBER CO., LTD.
? JANG, Ki-Ryong
? MOON, Jae-Woong
? BAE, Cheon-Soon
? YANG, Doo-Suk
? LEE, Jee-Won
? SEQNG, Baik-Lin
? TITLE OF INVENTION: Process for preparing recombinant proteins using highly
? efficient expression vector from Saccharomyces cerevisiae
? NUMBER OF SEQUENCES: 25
? CORRESPONDENCE ADDRESS:
? ADDRESSER: BACEMAN S LAPINTE, P.O.
? STREET: Suite 1261, 900 Chapel Street
? CITY: New Haven
? STATE: Connecticut
? COUNTRY: U.S.A.
? ZIP: 06510-2802
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
? COMPUTER: IBM
? OPERATING SYSTEM: WINDOWS 95/98
? SOFTWARE: MS WORD
? CURRENT APPLICATION DATA: US/09/424,6203
? FILING DATE: 24-MO. 5991595-1999
? INFORMATION FOR SEQ ID NO: 25:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 241 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: PROTEIN
? SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-424-620B-25

Query Match 97.9%; Score 460; DB 4; Length 241;
Best Local Similarity 98.9%; Pred. No. 8.2e-51;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 2 FTPIPLSLFDNMLRAHRLHQLAFDTYOEFEEAYIPKEQKYSFLONPQTSLSFSESPT 61
Db 51 FTPIPLSLFDNMLRAHRLHQLAFDTYOEFEEAYIPKEQKYSFLONPQTSLSFSESPT 110
QY 62 PSNREETQCKSNLELIRISLLLIQSWLEPVQ 92
Db 111 PSNREETQCKSNLELIRISLLLIQSWLEPVQ 141

RESULT 11
US-09-280-030-66
? Sequence 66, Application US/09280030A
? Patent No. 6506595
? GENERAL INFORMATION:
? APPLICANT: Sato, Seiji
? APPLICANT: Higashikuni, Naohiko
? APPLICANT: Kudo, Toshiyuki
? APPLICANT: Kondo, Masaki
? TITLE OF INVENTION: DNAS ENCODING NEW FUSION PROTEINS AND PROCESSES FOR
? TITLE OF INVENTION: PREPARING USEFUL POLYPEPTIDES THROUGH EXPRESSION OF THE
? TITLE OF INVENTION: DNAS
? FILE REFERENCE: 382.1026
? CURRENT APPLICATION NUMBER: US/09/280.030A
? CURRENT FILING DATE: 1999-03-26
? EARLIER APPLICATION NUMBER: JP10-87339/1998
? EARLIER FILING DATE: 1998-03-31
? NUMBER OF SEQ ID NOS: 66
? SOFTWARE: Patentin Ver. 2.0
? SEQ ID NO 66
? LENGTH: 245
? TYPE: PRT
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Description of Artificial Sequence: Designated is
? OTHER INFORMATION: an amino acid sequence of MWpMp20-TEV-G-GH
US-09-280-030-66

Query Match 97.9%; Score 460; DB 4; Length 245;
Best Local Similarity 98.9%; Pred. No. 8.4e-51;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FTPIPLSLFDNMLRAHRLHQLAFDTYOEFEEAYIPKEQKYSFLONPQTSLSFSESPT 61
Db 55 FTPIPLSLFDNMLRAHRLHQLAFDTYOEFEEAYIPKEQKYSFLONPQTSLSFSESPT 114
QY 62 PSNREETQCKSNLELIRISLLLIQSWLEPVQ 92
Db 115 PSNREETQCKSNLELIRISLLLIQSWLEPVQ 145

RESULT 12
US-08-784-582-71
? Sequence 71, Application US/08784582
? Patent No. 6110707
? GENERAL INFORMATION:
? APPLICANT: Newgard, Christopher H.
? APPLICANT: Halban, Philippe A.
? APPLICANT: No. 51079minston, Karl D.
? APPLICANT: Clark, Samuel A.
? APPLICANT: Thigpen, Anice E.
? APPLICANT: Quade, Christian
? APPLICANT: Kruse, Fred
? APPLICANT: McGarry, Dennis
? TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
? TITLE OF INVENTION: SECRETORY CELL LINES
? NUMBER OF SEQUENCES: 79
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Arnold, White & Durkee
? STREET: P.O. Box 4433
? CITY: Houston
? STATE: Texas
? COUNTRY: USA
? ZIP: 77210

```



```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/784,582
  FILING DATE: Concurrently Herewith
  CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 60/028,427
    FILING DATE: 15-OCT-1995
    NAME: Highlander, Steven L.
    REGISTRATION NUMBER: 37,642
    REFERENCE/DOCKET NUMBER: UTSD:514
    TELECOMMUNICATION INFORMATION:
      TELEPHONE: 512/418-3000
      TELEFAX: 512/474-7577
    INFORMATION FOR SEQ ID NO: 71:
      LENGTH: 254 amino acids
      TYPE: amino acid
      STRANDEDNESS:
      TOPOLOGY: linear
US-08-784-582-71

Query Match          97.9%; Score 460; DP 3; Length 274;
Best Local Similarity 98.9%; Pred. No. 9.9e-51;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FPTIPLSLFDNAMLRAHLHQLAFDTYQEFEEAYIPKECKYSFLONPOTSLSESISPT 61
DB 27 FPTIPLSLFDNAMLRAHLHQLAFDTYQEFEEAYIPKECKYSFLONPOTSLSESISPT 64

QY 62 PSNREETOQKSNLELLRISLLIQTOSWLEPVQ 92
DB 87 PSNREETOQKSNLELLRISLLIQTOSWLEPVQ 117

RESULT 13
US-08-784-582-73
Sequence 73, Application US/03784582
Patent No. 6110707
GENERAL INFORMATION:
  APPLICANT: Neward, Christopher B.
  APPLICANT: Halban, Philippe A.
  APPLICANT: No. 6110707/Minnton, Karl D.
  APPLICANT: Clark, Samuel A.
  APPLICANT: Thigpen, Anico E.
  APPLICANT: Quade, Christian
  APPLICANT: Kruse, Fred
  APPLICANT: McGarry, Dennis
  TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
  NUMBER OF SEQUENCES: 73
  CORRESPONDENCE ADDRESS:
    ADDRESS: Arnold, White & Durkee
    STREET: P.O. Box 4433
    CITY: Houston
    STATE: Texas
    COUNTRY: USA
    ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/784,582
```

```
FILING DATE: Concurrently Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 60/028,427
  FILING DATE: 15-OCT-1995
  NAME: Highlander, Steven L.
  REGISTRATION NUMBER: 37,642
  REFERENCE/DOCKET NUMBER: UTSD:514
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 512/418-3000
    TELEFAX: 512/474-7577
  INFORMATION FOR SEQ ID NO: 73:
    LENGTH: 350 amino acids
    TYPE: amino acid
    STRANDEDNESS:
    TOPOLOGY: linear
US-09-465-461-1

Query Match          97.9%; Score 450; DB 3; Length 360;
Best Local Similarity 98.9%; Pred. No. 1.4e-50;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FPTIPLSLFDNAMLRAHLHQLAFDTYQEFEEAYIPKECKYSFLONPOTSLSESISPT 61
DB 27 FPTIPLSLFDNAMLRAHLHQLAFDTYQEFEEAYIPKECKYSFLONPOTSLSESISPT 86

QY 62 PSNREETOQKSNLELLRISLLIQTOSWLEPVQ 92
DB 87 PSNREETOQKSNLELLRISLLIQTOSWLEPVQ 117

RESULT 14
US-09-465-461-1
Sequence 1, Application US/09465461
Patent No. 6348444
GENERAL INFORMATION:
  APPLICANT: CHAPPEL, Scott
  TITLE OF INVENTION: Human Growth Hormone to stimulate hematopoiesis and immune
  FILE REFERENCE: CHAPPEL=6;
  CURRENT APPLICATION NUMBER: US/09/465,461
  PRIOR FILING DATE: 1999-12-17
  PRIOR APPLICATION NUMBER: 60/112,668
  PRIOR FILING DATE: 1998-12-17
  NUMBER OF SEQ ID NOS: 1
  SOFTWARE: PatentIn version 3.1
  SEQ ID NO 1
  LENGTH: 191
  TYPE: PRT
  ORGANISM: homo sapiens
US-09-465-461-1

Query Match          96.6%; Score 454; DB 4; Length 191;
Best Local Similarity 97.8%; Pred. No. 3.5e-50;
Matches 89; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 FPTIPLSLFDNAMLRAHLHQLAFDTYQEFEEAYIPKECKYSFLONPOTSLSESISPT 61
DB 1 FPTIPLSLFDNAMLRAHLHQLAFDTYQEFEEAYIPKECKYSFLONPOTSLSESISPT 60

QY 62 PSNREETOQKSNLELLRISLLIQTOSWLEPVQ 92
DB 61 PSNREETOQKSNLELLRISLLIQTOSWLEPVQ 91

RESULT 15
US-08-187-756C-4
Sequence 4, Application US/08187756C
```

```

? Patent No. 5597769
? GENERAL INFORMATION:
? APPLICANT: ROSEN, ET AL
? TITLE OF INVENTION: Human Growth Hormone
? NUMBER OF SEQUENCES: 7
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: CARELIA, BYRNE, GAIN, SIMPILIAN,
? ADDRESS: GEOCHI, STEWART & CLISTHIN
? STREET: 6 BECKER FARM ROAD
? CITY: ROSELAND
? STATE: NEW JERSEY
? COUNTRY: USA
? ZIP: 07068
? COMPUTER READABLE FORM:
? MEDIUM TYPE: 3.5 INCH DISKETTE
? COMPUTER: IBM PS/2
? OPERATING SYSTEM: MS-DOS
? SOFTWARE: WORD PERFECT 5.1
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/187,756C
? FILING DATE: January 27, 1994
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER:
? FILING DATE:
? ATTORNEY/AGENT INFORMATION:
? NAME: FERRARO, GREGORY D.
? REGISTRATION NUMBER: 36,134
? REFERENCE/DOCKET NUMBER: 325800-55
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 201-994-1700
? TELEFAX: 201-594-1744
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 217 AMINO ACIDS
? TYPE: AMINO ACID
? STRANDEDNESS:
? TOPOLOGY: LINEAR
? MOLECULE TYPE: PROTEIN
? US-08-187-756C.4

Query Match 96.6% Score 454 Db 21 Length 217
Best Local Similarity 97.8% Pred. No. 4,20-50
Matches 89: Conservative G; Mismatches 2: Gaps 0

QY 2 PPTPLSLFLNANLPAHRLGLAFTYCEFEFRATPEKDYSELOKQCTSLSPSESIST 61
DB 22 PPTPLSLFLNANLPAHRLGLAFTYCEFEFRATPEKDYSELOKQCTSLSPSESIST 61
QY 62 PSNREETQCKSNLELRISLLICQSMLEPVC 52
DB 87 PSNREETQCKSNLELRISLLICQSMLEPVC 117

Search completed: September 15, 2003, 12:05:31
Job time : 22.1039 secs

```

GenCore version 3.1.6
Copyright (c) 1993-2003 Capgen Ltd.

CM protein - protein search, using sw mode)

Run on: September 15, 2003, 12:03:35 Search time 36,540.6 seconds
without alignment
547,945 Million cell of hits/sec

Title: US-09-423-100-2

Perfect score: 470

Sequence: 1 MEFTPLSRFDNMLKFAHR.....NLRKSTSLKLSGSLVIVD 42

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 541936 seqs, 14,912,426 residues

Total number of hits satisfying chosen parameters: 541936

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	470	100.0	92	16	US-10-054-873-2
2	470	100.0	150	13	US-10-054-873-7
3	450	97.9	191	13	US-09-984-010-23
4	463	97.9	191	12	US-10-153-207-1
5	460	97.9	191	13	US-10-400-377-1
6	460	97.9	191	12	US-10-400-708-1
7	460	97.9	214	12	US-10-153-207-6
8	460	97.9	217	9	US-09-929-318-9
9	450	97.9	245	9	US-09-280-030-56
10	454	96.6	217	9	US-09-853-688-2
11	454	96.6	217	14	US-09-943-7480-4
12	453	95.4	217	10	US-09-804-409A-16
13	447	95.1	217	9	US-09-853-688-4
14	444	94.5	191	12	US-09-824-200-12
15	396	84.3	217	9	US-09-850-887-3

381	61.1	153	15	US-10-043-487-350	Sequence 350, Appl
381	61.1	191	12	US-10-153-207-2	Sequence 2, Appl 1
373	73.4	229	15	US-10-103-313-41	Sequence 411, Appl
345	73.4	246	15	US-10-188-246-16	Sequence 18, Appl
301.5	64.1	190	12	US-10-153-207-3	Sequence 3, Appl 1
260	55.3	49	14	US-10-054-873-1	Sequence 1, Appl 1
260	55.3	107	14	US-10-054-873-6	Sequence 6, Appl 1
174.5	37.1	188	15	US-10-191-879-19	Sequence 19, Appl
174.5	37.1	210	15	US-10-191-879-10	Sequence 10, Appl
162	34.5	146	15	US-10-043-487-337	Sequence 337, Appl
161.5	34.4	366	10	US-09-887-569A-2	Sequence 2, Appl 1
161.5	34.4	388	12	US-10-322-746-4	Sequence 4, Appl 1
142	30.2	54	9	US-09-876-478-14	Sequence 14, Appl
132	27.9	196	15	US-10-140-293-32	Sequence 32, Appl
126	26.6	199	15	US-10-140-293-33	Sequence 33, Appl
125	26.6	199	15	US-10-140-293-25	Sequence 26, Appl
125	26.6	199	15	US-10-140-293-27	Sequence 27, Appl
124	26.4	199	15	US-10-140-293-31	Sequence 31, Appl
123	26.2	199	15	US-10-140-293-22	Sequence 22, Appl
122	26.0	199	15	US-10-140-293-21	Sequence 21, Appl
121.5	25.9	125	14	US-10-036-869-25	Sequence 25, Appl
121.5	25.9	206	15	US-10-191-879-2	Sequence 2, Appl 1
121.5	25.9	253	14	US-10-036-869-27	Sequence 27, Appl
119	25.3	199	15	US-10-140-293-24	Sequence 24, Appl
118	25.1	199	15	US-10-140-293-18	Sequence 18, Appl
116	25.1	199	15	US-10-140-293-19	Sequence 19, Appl
117	24.9	197	15	US-10-140-293-15	Sequence 15, Appl
116	24.7	197	15	US-10-140-293-13	Sequence 13, Appl
116	24.7	199	15	US-10-153-207-4	Sequence 4, Appl 1
116	24.7	199	15	US-10-140-293-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-10-054-873-2
Sequence 2, Application US/10054873
Publication No. US2002016472A1
GENERAL INFORMATION:
APPLICANT: Gan, Zhong Ru
TITLE OF INVENTION: Chimeric Protein Containing an Intramolecular Chaperone-Like Sequence
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/10/054,873
APPLICATION NUMBER: US/10/054,873
FILING DATE: 22-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/CN98/00052
FILING DATE: 31-MAR-1998
APPLICATION NUMBER: US 09/423,100
FILING DATE: 11-DEC-2000
ATTORNEY/AGENT INFORMATION:
NAME: Mycroft, Frank J.
REGISTRATION NUMBER: 46,946
REFERENCE/DOCKET NUMBER: 020167-000130US
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 92 amino acids
TYPE: amino acid

```

: STRANDEDNESS: <Unknown>
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-054-873-2

Query Match      100.0%; Score 470; DB 14; Length 150;
Best Local Similarity 100.0%; Pred. No. 2.2e-46;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPTTPLSRFNFAMLAHRLHQLAFDYOEFEAYIPKEQKYSFLQNPQTSLSSESIP 60
DB 1 MPTTPLSRFNFAMLAHRLHQLAFDYOEFEAYIPKEQKYSFLQNPQTSLSSESIP 60
QY 61 TPSNRETOOKSNLELRISLLIOSWLEPVQ 92
DB 61 TPSNRETOOKSNLELRISLLIOSWLEPVQ 92

RESULT 2
US-10-054-873-7
: Sequence 7, Application US/10054873
: Publication No. US20020164712A1
: GENERAL INFORMATION:
: APPLICANT: Gan, Zhong Rq
: TITLE OF INVENTION: Chimero Protein Containing an
: Intramolecular Chaperone Like Sequence
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Townsend and Crew LLP
: STREET: Two Embarcadero Center, Eighth Floor
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-3834
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/10/054,873
: FILING DATE: 22-Jan-2002
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: WO PCT/CN98/000352
: FILING DATE: 31-MAR-1998
: APPLICATION NUMBER: US 09/423,100
: FILING DATE: 21-DEC-2000
: ATTORNEY/AGENT INFORMATION:
: NAME: Mycroft, Frank J
: REGISTRATION NUMBER: 44,946
: REFERENCE/DOCKET NUMBER: 020167-0001300S
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 150 amino acids
: TYPE: amino acid
: STRANDEDNESS: <Unknown>
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-054-873-7

Query Match      100.0%; Score 470; DB 14; Length 150;
Best Local Similarity 100.0%; Pred. No. 4.1e-46;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPTTPLSRFNFAMLAHRLHQLAFDYOEFEAYIPKEQKYSFLQNPQTSLSSESIP 60
DB 1 MPTTPLSRFNFAMLAHRLHQLAFDYOEFEAYIPKEQKYSFLQNPQTSLSSESIP 60
QY 61 TPSNRETOOKSNLELRISLLIOSWLEPVQ 92
DB 61 TPSNRETOOKSNLELRISLLIOSWLEPVQ 92
```

```

DB 61 TPSNRETOOKSNLELRISLLIOSWLEPVQ 92

RESULT 3
US-09-984-010-23
: Sequence 23, Application US/09984010
: Publication No. US20030104578A1
: GENERAL INFORMATION:
: APPLICANT: Ballance, David James
: TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
: NUMBER OF SEQUENCES: 26
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP
: STREET: 1300 I Street, NW
: CITY: Washington
: STATE: DC
: COUNTRY: USA
: ZIP: 20005-3315
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/984,010
: FILING DATE: 21-May-2002
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 09/091,873
: FILING DATE: 25-JUN-1998
: APPLICATION NUMBER: PCT/GR96/03164
: FILING DATE: 19-DEC-1996
: INFORMATION FOR SEQ ID NO: 23:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 191 amino acids
: TYPE: amino acid
: STRANDEDNESS: <Unknown>
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHEICAL: NO
: ANTI-SENSE: NO
: SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-984-010-23

Query Match      97.9%; Score 460; DB 11; Length 191;
Best Local Similarity 98.9%; Pred. No. 7.8e-45;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PPTPLSRFNFAMLAHRLHQLAFDYOEFEAYIPKEQKYSFLQNPQTSLSSESIP 61
DB 1 PPTPLSRFNFAMLAHRLHQLAFDYOEFEAYIPKEQKYSFLQNPQTSLSSESIP 60
QY 62 PSNRETOOKSNLELRISLLIOSWLEPVQ 92
DB 61 PSNRETOOKSNLELRISLLIOSWLEPVQ 91

RESULT 4
US-10-153-207-1
: Sequence 1, Application US/10153207
: Publication No. US20030153003A1
: GENERAL INFORMATION:
: APPLICANT: James A. Wells
: APPLICANT: Brian C. Cunningham
: TITLE OF INVENTION: GROWTH HORMONE VARIANTS
: FILE REFERENCE: 669.12-US-C7
: CURRENT APPLICATION NUMBER: US/10/153,207
: CURRENT FILING DATE: 2002-05-22
: PRIOR APPLICATION NUMBER: 08/479,884
: PRIOR FILING DATE: 1995-06-07
: PRIOR APPLICATION NUMBER: 08/190,723
: PRIOR FILING DATE: 1994-02-02
: PRIOR APPLICATION NUMBER: 07/960,227
```

```

; PRIOR FILING DATE: 1992-10-13
; PRIOR APPLICATION NUMBER: 07/625,204
; PRIOR FILING DATE: 1992-04-27
; PRIOR APPLICATION NUMBER: 07/424,066
; PRIOR FILING DATE: 1989-10-25
; PRIOR APPLICATION NUMBER: 07/264,611
; PRIOR FILING DATE: 1988-10-23
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 1
; LENGTH: 191
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-400-708-1

Query Match          97.9%   Score 460; DB 12; Length 191;
Best Local Similarity 98.9%   Pred. No. 7.8e-45;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FPTPLSRFLPDNAMLRAHRLHQLAFDTYQEFEEAYIPKEOKYSFLQNPTSLSFSES IPT 61
DB 1 FPTPLSRFLPDNAMLRAHRLHQLAFDTYQEFEEAYIPKEOKYSFLQNPTSLSFSES IPT 60

QY 62 PSNREETOQKSNLELLKISLLLIQSWLEPVQ 92
DB 61 PSNREETOQKSNLELLKISLLLIQSWLEPVQ 91

RESULT 5
US-10-400-377-1
; Sequence 1, Application US/10400377
; Publication No. US20030162949A1
; GENERAL INFORMATION:
; APPLICANT: Cox III, George N
; TITLE OF INVENTION: Derivatives of Growth Hormone and Related Proteins
; FILE REFERENCE: 4:52-1-PUS
; CURRENT APPLICATION NUMBER: US/10/400,377
; CURRENT FILING DATE: 2003-03-25
; PRIOR APPLICATION NUMBER: US/09/462,941
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 60/052,516
; PRIOR FILING DATE: 1997-07-14
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 1
; LENGTH: 191
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-400-377-1

Query Match          97.9%   Score 460; DB 12; Length 191;
Best Local Similarity 98.9%   Pred. No. 7.8e-45;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FPTPLSRFLPDNAMLRAHRLHQLAFDTYQEFEEAYIPKEOKYSFLQNPTSLSFSES IPT 61
DB 1 FPTPLSRFLPDNAMLRAHRLHQLAFDTYQEFEEAYIPKEOKYSFLQNPTSLSFSES IPT 60

QY 62 PSNREETOQKSNLELLKISLLLIQSWLEPVQ 92
DB 61 PSNREETOQKSNLELLKISLLLIQSWLEPVQ 91

RESULT 6
US-10-400-708-1
; Sequence 1, Application US/10400708
; Publication No. US20030166865A1
; GENERAL INFORMATION:
; APPLICANT: Cox III, George N
; TITLE OF INVENTION: Derivatives of Growth Hormone and Related Proteins
; FILE REFERENCE: 4:52-1-PUS

```

```

; CURRENT APPLICATION NUMBER: US/10/400,708
; CURRENT FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: US/09/462,941
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 60/052,516
; PRIOR FILING DATE: 1997-07-14
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 1
; LENGTH: 191
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-400-708-1

Query Match          97.9%   Score 460; DB 12; Length 191;
Best Local Similarity 98.9%   Pred. No. 7.8e-45;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FPTPLSRFLPDNAMLRAHRLHQLAFDTYQEFEEAYIPKEOKYSFLQNPTSLSFSES IPT 61
DB 1 FPTPLSRFLPDNAMLRAHRLHQLAFDTYQEFEEAYIPKEOKYSFLQNPTSLSFSES IPT 60

QY 62 PSNREETOQKSNLELLKISLLLIQSWLEPVQ 92
DB 61 PSNREETOQKSNLELLKISLLLIQSWLEPVQ 91

RESULT 7
US-10-153-207-6
; Sequence 6, Application US/10153207
; Publication No. US2003015303A1
; GENERAL INFORMATION:
; APPLICANT: James A. Wells
; APPLICANT: Brian C. Cunningham
; TITLE OF INVENTION: GROWTH HORMONE VARIANTS
; FILE REFERENCE: 669.12-US-C7
; CURRENT APPLICATION NUMBER: US/10/153,207
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: 08/479,884
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 08/190,723
; PRIOR FILING DATE: 1994-02-02
; PRIOR APPLICATION NUMBER: 07/960,227
; PRIOR FILING DATE: 1992-10-13
; PRIOR APPLICATION NUMBER: 07/875,204
; PRIOR FILING DATE: 1992-04-27
; PRIOR APPLICATION NUMBER: 07/428,066
; PRIOR FILING DATE: 1989-10-26
; PRIOR APPLICATION NUMBER: 07/264,611
; PRIOR FILING DATE: 1988-10-28
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 6
; LENGTH: 214
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-153-207-6

Query Match          97.9%   Score 460; DB 12; Length 214;
Best Local Similarity 98.9%   Pred. No. 9e-45;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FPTPLSRFLPDNAMLRAHRLHQLAFDTYQEFEEAYIPKEOKYSFLQNPTSLSFSES IPT 61
DB 24 FPTPLSRFLPDNAMLRAHRLHQLAFDTYQEFEEAYIPKEOKYSFLQNPTSLSFSES IPT 83

QY 62 PSNREETOQKSNLELLKISLLLIQSWLEPVQ 92
DB 84 PSNREETOQKSNLELLKISLLLIQSWLEPVQ 114

RESULT 8
US-09-929-918-9

```



```

1 REFERENCE/DOCKET NUMBER: PF-0535 US
2
3 TELECOMMUNICATION INFORMATION:
4 TELEPHONE: (650) 855-0555
5 TELEFAX: (650) 845-4166
6
7 INFORMATION FOR SEQ ID NO: 1:
8     SEQUENCE CHARACTERISTICS:
9         LENGTH: 217 amino acids
10        TYPE: amino acid
11        STRANDEDNESS: single
12        TOPOLOGY: linear
13
14 IMMEDIATE SOURCE:
15     LIBRARY: GenBank
16     CLONE: c404987
17
18     SEQUENCE DESCRIPTION: SEQ ID NO: 1 :
19
20 US-09-650-887-5
21
22 Query Match      84.3%   Score 396, DB 9;   Length 217;
23 Best Local Similarity 84.6%,   Prod. No. 1,9e-37;
24 Matches 77; Conservative 6; Mismatches 8; Indels 0; Gaps 0;
25
26 QY 2 FPTITLKRLEPDNAMIRHRLRQLAQITQYGFEEATPKPKQKYSPLQNTQTSLSSESPT 41
27      111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
28 DB 27 FPTITPLSWLNTAVFRAHHKRLKLAFTYKPKLEAAVFKKQKYSPLQNTQTSLSSESPT 86
29      111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
30
31 QY 62 PSNKEETQCKSNLELRISLSTQSWLPEVQ 92
32      111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
33 DB 87 PSNKEETQCKSNLELRISLSTQSWLPEVQ 117
34      111 111 111 111 111 111 111 111 111 111 111 111 111 111 111

```

Search completed: September 15, 2003, 12:23:15
Job time : 39.5806 secs

A;Cross-references: GB:M14398; NED:q163158; PIDN:AAA52554.1; PID:q163159

267409

267409

A:Molecule type: protein
A:Residues: 27-217 <NIA>
A:Experimental source: placenta
B:McC A Baird, N.; Tiplon, K.F
Biochem. Soc. Trans. 19, 205, 1991
A:Title: Catechol-O-methyltransferase from human placenta: purification and some properties
A:Reference number: A51283; MUID:91244006; PMID:2947148
A:Accession: A61283
A:Molecule type: protein
A:Residues: 27-46 <NIC>
A:Note: chorionamniotropon apparently copurified with placenta. Catechol-O-methyltransferase
R:Sherwood, L.M.; Handwerker, S.; McCaurin, W.B.; Lander, M.
Nature New Biol. 233, 59-62, 1971
A:Title: Amino-acid sequence of human placental catechol-O-methyltransferase
A:Reference number: A93431; MUID:72016111; PMID:5246343
A:Contents: annotation
R:Sherwood, L.M.; Handwerker, S.; McCaurin, W.B.; Lander, M.
Nature New Biol. 233, 54, 1972
A:Reference number: A93405
A:Contents: annotation
R:Schneller, A.H.; Kowalski, K.; Russell, J.; Sherwood, L.M.
J. Biol. Chem. 254, 3782-3787, 1979
A:Title: Identification of the interchain disulfide bonds of human chorionamniotropon. In
A:Reference number: A92251; MUID:79173081; PMID:746153
A:Contents: annotation; dimeric disulfide bonds
R:Selby, M.J.; Barta, A.; Baxter, J.D.; Bell, G.F.; Eberhardt, N.L.
J. Biol. Chem. 259, 13131-13138, 1984
A:Title: Analysis of a major human chorionic somatomamotropin gene. Evidence for two
A:Reference number: 155229; MUID:85030425; PMID:6288192
A:Accession: 155229
A:Status: translated from GB/EMBL/DDA
A:Molecule type: DNA
A:Residues: 1-217 <RES>
A:Cross-references: GB:K02401; NID:q181120; PIDN:AAA52115.1; EMBL:q181121
R:Seeburg, P.H.; Shine, J.; Martial, J.A.; Ullrich, A.; Goodman, H.
Trans. Assoc. Am. Physicians 90, 103-116, 1977
A:Title: Nucleotide sequence of a human gene coding for a polypeptide hormone.
A:Reference number: 159658; MUID:78160787; PMID:611657
A:Accession: 159658
A:Status: translated from GB/EMBL/DDA
A:Molecule type: mRNA
A:Residues: 150-217 <RE2>
A:Cross-references: GB:M25118; NID:q181124; PIDN:AAA35721.1; PMID:q181123
C:Genetics:
A:Gene: GDB:CSH1
A:Cross-references: GDB:119084; OMIM:150200
A:Map position: 17q22-17q24
A:Introns: 4/1: 57/3: 97/3: 152/3
C:Superfamily: prolactin
C:Keywords: hormone; placenta
F:1-26/Domain: signal sequence #status experimental <SIG>
F:27-217/Product: chorionamniotropon #status experimental <MAT>
F:79-191/Disulfide bonds: #status experimental
F:208-215/Disulfide bonds: (in monomeric form) #status experimental
F:208/Disulfide bonds: interchain (to 215 in dimeric form) #status experimental
F:215/Disulfide bonds: interchain (to 208 in dimeric form) #status experimental

Query Match 81.1% Score 381; DB 1; Length 217;
Best Local Similarity 82.0% Pred. No. 1.4e-33;
Matches 73; Conservative 8; Mismatches 0; Indels 0; Gaps 0;
QY 4 TIPSRLFDNAMLRAHRLHQLADTYOEFEEYIPKEQKYSFLQNPTSLSESIPTVS 63
DB 29 TVPLSRFDHMLQAHRAHQLADTYOEFETYPKQKYSFLHDSQISFCSDSIPTVS 88
QY 64 NREETOQKSNLELLRLISLLIQSWELEPV 92
DB 89 NMEETOQKSNLELLRLISLLIQSWELEPV 117

RESULT 10
E32435
chorionamniotropon H precursor - human

N:Alternate names: chorionic somatomamotropin 2
C:Species: Homo sapiens (man)
C:Date: 29-Dec-1989 #sequence_revision 29-Dec-1989 #text_change 16-Jul-1999
C:Accession: E32435
R:Chen, E.Y.; Liao, Y.C.; Smith, D.H.; Herrera-Saldana, H.A.; Gelinas, R.E.; Seeburg, P.H.; Shine, J.; Martial, J.A.; Ullrich, A.; Goodman, H.
Genomics 4, 479-497, 1989
A:Title: The human growth hormone locus: nucleotide sequence, biology, and evolution
A:Reference number: A32435; MUID:89307277; PMID:2744760
A:Accession: E32435
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-217 <CHE>
A:Cross-references: GB:JC3071; NID:q183148; PIDN:AAA52553.1; PID:q183153
C:Genetics:
A:Gene: GDB:CSH2
A:Cross-references: GDB:119813; OMIM:118820
A:Map position: 17q22-17q24
C:Superfamily: prolactin

Query Match 81.1% Score 381; DB 2; Length 217;
Best Local Similarity 82.0% Pred. No. 1.4e-33;
Matches 73; Conservative 8; Mismatches 0; Indels 0; Gaps 0;
QY 4 TIPSRLFDNAMLRAHRLHQLADTYOEFEEYIPKEQKYSFLQNPTSLSESIPTVS 63
DB 29 TVPLSRFDHMLQAHRAHQLADTYOEFETYPKQKYSFLHDSQISFCSDSIPTVS 88
QY 64 NREETOQKSNLELLRLISLLIQSWELEPV 92
DB 89 NMEETOQKSNLELLRLISLLIQSWELEPV 117

RESULT 11
A26449
chorionamniotropon precursor (allele hCS-3) - human
C:Species: Homo sapiens (man)
C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 28-Jul-1995
C:Accession: A26449
R:Birt, H.; Kimmel, J.; Birnbaum, J.; Chen, E.Y.; Seeburg, P.H.; Eberhardt, N.L.
DNA 6, 59-70, 1987
A:Title: The human growth hormone gene locus: structure, evolution, and allelic variation
A:Reference number: A26449; MUID:87161235; PMID:3030680
A:Accession: A26449
A:Molecule type: DNA
A:Residues: 1-215 <HIR>
C:Superfamily: prolactin
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-215/Product: chorionamniotropon, hCS-3 allele #status predicted <MAT>

Query Match 76.5% Score 359.5; DB 2; Length 215;
Best Local Similarity 80.5% Pred. No. 2.9e-31;
Matches 70; Conservative 8; Mismatches 1; Gaps 1;
QY 4 TIPSRLFDNAMLRAHRLHQLADTYOEFEEYIPKEQKYSFLQNPTSLSESIPTVS 63
DB 29 TVPLSRFDHMLQAHRAHQLADTYOEFETYPKQKYSFLHDSQISFCSDSIPTVS 88
QY 64 NREETOQKSNLELLRLISLLIQSWELEPV 90
DB 89 NMEETOQKSNLELLRLISLLIQSWELEPV 114

RESULT 12
A49159
somatotropon - golden hamster
N:Alternate names: growth hormone
C:Species: Mesocricetus auratus (golden hamster)
C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C:Accession: B49159
R:Southard, J.N.; Sanchez-Jimenez, F.; Campbell, G.T.; Tolamantes, F.
Endocrinology 129, 2965-2971, 1991
A:Title: Sequence and expression of hamster prolactin and growth hormone messenger
A:Reference number: A49159; MUID:92063850; PMID:1954881

GenInfo version 5.1.1
Copyright (c) 1993 - 2003 CompuLink Ltd.

EM protein - protein search, using sw model

Run on: September 15, 2003, 11:54:06 : Search time 0:17:48 seconds
(without alignment)
594.63 : BLAST2S Cell: updates/seq

Title: US-09-423-100-2

Perfect score: 470

Sequence: 1 MFPIPLSKFENAMRAHP.....NFTLRNLSLISNIMRNVQ 74

Scoring table: BLOSUM62

Gapop 16.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 6

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 60

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.7

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	460	97.9	217	1	SOMA_HUMAN
2	460	97.9	217	1	SOMA_MACMU
3	460	97.9	217	1	SOMA_PANTR
4	433	92.3	217	1	SOMA_SALIB
5	432	91.9	217	1	SOMA_GALLA
6	430	91.5	217	1	SOMA_PANTR
7	422	89.8	217	1	SOMA_HUMAN
8	396	84.3	217	1	SOMA_MACMU
9	391	81.1	217	1	PLC_HUMAN
10	310.5	66.1	215	1	SOMA_MESAU
11	307.5	65.4	190	1	SOMA_BALBO
12	304.5	64.8	216	1	SOMA_MOUSE
13	302.5	64.4	216	1	SOMA_HORSE
14	302.5	64.4	216	1	SOMA_RABIT
15	302.5	64.4	216	1	SOMA_RAT
16	302.5	64.4	217	1	SOMA_GALSE
17	302.5	64.4	217	1	SOMA_NYCPV
18	301.5	64.1	190	1	SOMA_GALAX
19	301.5	64.1	216	1	SOMA_PANPA
20	301.5	64.1	216	1	SOMA_FELCA
21	301.5	64.1	216	1	SOMA_PYG
22	299.5	63.7	215	1	SOMA_MOUSE
23	297.5	63.3	190	1	SOMA_LAKKA
24	295.5	62.9	190	1	SOMA_VULVU
25	291.5	62.0	215	1	SOMA_MONIC
26	291.5	62.0	215	1	SOMA_TKIVU
27	289.5	61.6	217	1	SOMA_KOVIN
28	289.5	61.6	217	1	SOMA_OBEL
29	289.5	61.6	217	1	SOMA_SHEEP
30	282.5	60.1	217	1	SOMA_RABBO
31	278.5	59.3	215	1	SOMA_MELGA
32	275.5	58.6	216	1	SOMA_CHICK
33	274.5	58.4	217	1	SOMA_STRICA

34 272.5 58.0 190 1 SOMA_CRONO
35 268.5 57.1 191 1 SOMA_CHEMY
36 261 55.5 216 1 SOMA_ANAPI
37 257.5 54.8 190 1 SOM1_ACIGU
38 257.5 54.8 190 1 SOM2_ACIGU
39 247.5 52.7 215 1 SOMA_LEPOS
40 239.5 51.0 214 1 SOMA_XENIA
41 238.5 50.7 215 1 SOMA_KANCA
42 226.5 48.2 215 1 SOMA_BUFMA
43 225.5 48.0 183 1 SOMA_PRIGL
44 219.5 46.7 208 1 SOMB_XENIA
45 218.5 46.5 206 1 SOMA_PROAN

10 SOMA_HUMAN STANDARD: PRT: 217 AA.
AC F01241: Q14405; Q16631; Q9H521; Q9UMJ7; Q9UNC5;
BT 21-JUL-1986 (Rel. 01. created)
BT 01-MAR-1992 (Rel. 21. last sequence update)
BT 15-SEP-2003 (Rel. 42. last annotation update)
CE Somatotropin precursor (Growth hormone) (GH) (GH-N) (pituitary growth hormone) (Growth hormone 1).
DE Hormone
GN GH1
OS Homo sapiens (Human).
SC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
GX NCBI_TaxID=9606;
RN 1
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=80034477; PubMed=386281;
RA Roskam W., Rougeon F.
RT "Molecular cloning and nucleotide sequence of the human growth hormone structural gene."
RI Hormone structural gene.
RL Nucleic Acids Res. 7:305-320(1979).
RN 12
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=79203293; PubMed=377496;
RA Martini J.A., Hallewell R.A., Baxter J.D., Goodman H.M.;
RT "Human growth hormone: complementary DNA cloning and expression in bacteria."
RL Science 205:602-607(1979).
RN 13
RP SEQUENCE FROM N.A. (ISOFORM 1), AND POSSIBLE ALTERNATIVE SPLICING.
RX MEDLINE=82014939; PubMed=6269091;
RA Denoto F.M., Moore D.D., Goodman H.M.;
RT "Human growth hormone DNA sequence and mRNA structure: possible alternative splicing."
RI Nucleic Acids Res. 9:3719-3730(1981).
RN 14
RP SEQUENCE FROM N.A.
RX MEDLINE=83182010; PubMed=7169009;
RA Seeburg P.H.;
RT "The human growth hormone gene family: nucleotide sequences show recent divergence and predict a new polypeptide hormone."
RI DNA 1:239-249(1982).
RN 15
RP SEQUENCE FROM N.A.
RX MEDLINE=89407277; PubMed=2744760;
RA Chen E.Y., Liao Y.C., Smith D.H., Barrera-Saldana H.A.,
RT "The human growth hormone locus: nucleotide sequence, biology, and evolution."
RI Genomics 4:479-457(1989).
RN 16
RP SEQUENCE FROM N.A. (ISOFORM 3).
RX TISSUE=Pituitary;
RA Gu G., Huang Q.-H., Li N., Xu S.-H., Han Z.-G., Fu G., Chen Z.;
RT "A novel gene expressed in human pituitary."
RI Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

ALIGNMENTS

RESULT 1

SOMA_HUMAN

10 SOMA_HUMAN STANDARD: PRT: 217 AA.

AC F01241: Q14405; Q16631; Q9H521; Q9UMJ7; Q9UNC5;

BT 21-JUL-1986 (Rel. 01. created)

BT 01-MAR-1992 (Rel. 21. last sequence update)

BT 15-SEP-2003 (Rel. 42. last annotation update)

CE Somatotropin precursor (Growth hormone) (GH) (GH-N) (pituitary growth hormone) (Growth hormone 1).

DE Hormone

GN GH1

OS Homo sapiens (Human).

SC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

GX NCBI_TaxID=9606;

RN 1

RP SEQUENCE FROM N.A. (ISOFORM 1).

RX MEDLINE=80034477; PubMed=386281;

RA Roskam W., Rougeon F.;

RT "Molecular cloning and nucleotide sequence of the human growth hormone structural gene."

RI Hormone structural gene.

RL Nucleic Acids Res. 7:305-320(1979).

RN 12

RP SEQUENCE FROM N.A. (ISOFORM 1).

RX MEDLINE=79203293; PubMed=377496;

RA Martini J.A., Hallewell R.A., Baxter J.D., Goodman H.M.;

RT "Human growth hormone: complementary DNA cloning and expression in bacteria."

RL Science 205:602-607(1979).

RN 13

RP SEQUENCE FROM N.A. (ISOFORM 1), AND POSSIBLE ALTERNATIVE SPLICING.

RX MEDLINE=82014939; PubMed=6269091;

RA Denoto F.M., Moore D.D., Goodman H.M.;

RT "Human growth hormone DNA sequence and mRNA structure: possible alternative splicing."

RI Nucleic Acids Res. 9:3719-3730(1981).

RN 14

RP SEQUENCE FROM N.A.

RX MEDLINE=83182010; PubMed=7169009;

RA Seeburg P.H.;

RT "The human growth hormone gene family: nucleotide sequences show recent divergence and predict a new polypeptide hormone."

RI DNA 1:239-249(1982).

RN 15

RP SEQUENCE FROM N.A.

RX MEDLINE=89407277; PubMed=2744760;

RA Chen E.Y., Liao Y.C., Smith D.H., Barrera-Saldana H.A.,

RT "The human growth hormone locus: nucleotide sequence, biology, and evolution."

RI Genomics 4:479-457(1989).

RN 16

RP SEQUENCE FROM N.A. (ISOFORM 3).

RX TISSUE=Pituitary;

RA Gu G., Huang Q.-H., Li N., Xu S.-H., Han Z.-G., Fu G., Chen Z.;

RT "A novel gene expressed in human pituitary."

RI Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.


```

CC use by non-profit institutions as long as the content is not modified and this statement is not removed. Usage by other parties requires a license agreement. (See http://www.isb-sib.ch/about) or send an email to license@sib-sib.ch)
CC -----
DR EMBL: K00470: AAA38612.1;
DR EMBL: J03756: AAB59547.1;
DR EMBL: J03756: AAB59548.1;
DR EMBL: M38451: AAA35891.1;
DR EMBL: C03071: AAA52552.1;
DR EMBL: BC020760: AAH20760.1;
DR PIR: A28072: SHUV2;
DR PIR: D32435: SHUV;
DR HSP: P01241: IAZ2;
DR Genew: HGNC:4262: GH2;
DR MIM: 139240;
DR GO: 0005180: F:peptide hormone; TAS;
DR InterPro: IPR001400: Somatotropin;
DR Pram: PF00103: hormone; 1;
DR PROSITE: PS00266: SOMATOTROPIN_1; 1;
DR PROSITE: PS00339: SOMATOTROPIN_2; 1;
KW Hormone; Placenta; Signal; Glycoprotein; Alternative splicing;
KW Polymorphism;
FT SIGNAL 1 26
FT CHAIN 27 217
FT DISULFID 79 191
FT DISULFID 208 215
FT CARBOHYD 166 166
FT VARSPLIC 153 217
FT VARIANT 90 90
FT VARIANT 109 109
FT VARIANT 169 169
FT VARIANT 217 AA: 24999 MW: 799324698E22F96 CRC64:
FT CONFLICT 109 109 I -> T (IN REF. 2)
FT CONFLICT 217 AA: 24999 MW: 799324698E22F96 CRC64:
SQ SEQUENCE 217 AA: 24999 MW: 799324698E22F96 CRC64:
Query Match 89.8%; Score 422; DB 1; Length 217;
Best Local Similarity 92.3%; Pred. No. 7.8e-37;
Matches 84; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
CY 2 FPIPLSRFDNMLRAHRLHQLAFDTYQFFRATYTPKQKYSFLONPOTSLSESSEIPT 61
Db 27 FPIPLSRFDNMLRAHRLHQLAFDTYQFFRATYTPKQKYSFLONPOTSLSESSEIPT 66
CY 62 PSNRRETCQKSNLELRISLLIQSWLSPVQ 92
Db 87 PSNRVTCQKSNLELRISLLIQSWLSPVQ 117
RESULT 8
SOM2_MACMU STANDARD; PRT; 217 AA.
AC Q07370: Q28494;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Growth hormone variant precursor (GH-V) (placenta-specific growth hormone) (growth hormone 2);
GN GH2;
OS Macaca mulatta (Rhesus macaque);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Macaca;
OC Cercopithecoidea; Macaca;
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA Golos T.G.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.

```

```

TISSE-Placenta;
MEDLINE=94048724; PubMed=4434617;
RA Golos T.G., Durning M., Fisher J.M., Fowler P.D.;
RT Cloning of four growth hormone/chorionic somatomammotropin-related complementary deoxyribonucleic acids differentially expressed during pregnancy in the rhesus monkey placenta;
RL Endocrinology 133:1744-1752(1993);
CC -!- FUNCTION: Plays an important role in growth control. Its major role in stimulating body growth is to stimulate the liver and other tissues to secrete IGF-1. It stimulates both the differentiation and proliferation of myoblasts. It also stimulates amino acid uptake and protein synthesis in muscle and other tissues.
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- TISSUE SPECIFICITY: Expressed in the placenta.
CC -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/about) or send an email to license@sib-sib.ch)
CC -----
DR EMBL: J03756: AAA38612.1;
DR EMBL: J03756: AAB59547.1;
DR PIR: 167411: 167411;
DR HSP: P01241: IHC9;
DR InterPro: IPR001400: Somatotropin;
DR Pram: PF00103: hormone; 1;
DR PROSITE: PS00266: SOMATOTROPIN_1; 1;
DR PROSITE: PS00338: SOMATOTROPIN_2; 1;
KW Hormone; Placenta; Signal; Glycoprotein;
FT SIGNAL 1 26
FT CHAIN 27 217
FT DISULFID 79 191
FT DISULFID 208 215
FT CONFLICT 57 57 L -> F (IN REF. 2);
FT CONFLICT 152 152 E -> G (IN REF. 2);
SQ SEQUENCE 217 AA: 25221 MW: 8DB166BC24EA09C CRC64:
Query Match 84.3%; Score 396; DB 1; Length 217;
Best Local Similarity 94.6%; Pred. No. 3.9e-34;
Matches 77; Conservative 6; Mismatches 8; Indels 0; Gaps 0;
CY 2 FPIPLSRFDNMLRAHRLHQLAFDTYQFFRATYTPKQKYSFLONPOTSLSESSEIPT 61
Db 27 FPIPLSRFDNMLRAHRLHQLAFDTYQFFRATYTPKQKYSFLONPOTSLSESSEIPT 86
CY 62 PSNRRETCQKSNLELRISLLIQSWLSPVQ 92
Db 87 PSNRVTCQKSNLELRISLLIQSWLSPVQ 117
RESULT 9
SOM2_HUMAN STANDARD; PRT; 217 AA.
AC P01243;
DT 21-JUL-1985 (Rel. 01, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Lactogen precursor (Chorion somatomotropin) (chorionic somatomammotropin). CSH1 AND CSH2.
GN CSH1 AND CSH2;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (GENE CSH1)
RA Selby M.J., Barta A., Baxter J.D., Bell G.I., Eberhardt N.L.;

```

RT "Analysis of a major human chorionic somatomammotropin gene. Evidence for two functional promoter elements." J. Biol. Chem. 259:13131-13138(1984).

RN [2]

RP SEQUENCE FROM N.A. (GENE CSH2).

RX MEDLINE=87161235; PubMed=303580;

RA Hift H., Kime-man J., Birnbaum M.J., Chen E.Y., Seeburg P.H., Eberhardt N.E., Bara A.

RT "The human growth hormone gene locus: structure, evolution, and allelic variations."

RL DNA 6:59-70(1987).

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=83163916; PubMed=6300056;

RA Barrera-Salgada H.A., Secoy P.H., Saunders G.F.

RT "Two structurally different genes produce the same placental human placental lactogen hormone."

RL J. Biol. Chem. 259:3787-3793(1984).

RN [4]

RP SEQUENCE FROM N.A. (GENES CSH1 AND CSH2).

RX MEDLINE=83307277; PubMed=2744769;

RA Chen E.Y., Liao Y.C., Smith D.H., Barrera-Salgada H.A., Gelinas R.E., Seeburg P.H.

RT "The human growth hormone locus: nucleotide sequence, biology, and evolution."

RL Genomics 4:479-497(1989).

RN [5]

RP SEQUENCE.

RX MEDLINE=83182610; PubMed=7169069;

RA Seeburg P.H.

RT "The human growth hormone gene family: nucleotide sequences show recent divergence and predict a new polypeptide hormone."

RL DNA 1:239-243(1982).

RN [6]

RP SEQUENCE FROM N.A.

RC TISSUE=Placenta, and Uterus;

RX MEDLINE=22386257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.E., Berger J.S., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Scheller G.D., Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P., Dudchenko L., Karusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., McLachy S., Bosak S.A., McBean P.J., McKernan K.J., Malek J.A., Gunaratne P., Richards S., Worley K.C., Hale S., Garcia A.M., Gay D., Huiyk S.W., Villalon D.K., Muzny D.M., Sodergren E., Litz J., Gibbs R.A., Fahey J., Helton E., Kettner A., Madan A., Rochevsky V., Bouffard A., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Battersfield V.S.N., Kravitz M.J., Skalska D., Smolnik D.E., Schnerch A., Schein J.E., Jones S.J.M., Mitra N.A., et al.

RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."

RL Proc. Natl. Acad. Sci. U.S.A. 99:15899-15903(2002).

RN [7]

RP SEQUENCE OF 50-217 FROM N.A.

RX MEDLINE=78071761; PubMed=553368;

RA Shiao J., Seeburg P.H., Martial C.A., Baxter J.D., Goodman R.E.

RT "Construction and analysis of recombinant DNA for human chorionic somatomammotropin."

RL Nature 270:494-499(1977).

RN [8]

RP SEQUENCE OF 27-217.

RX MEDLINE=73201971; PubMed=4712450;

RA Li C.H., Dixon J.S., Chang D.

RT "Amino acid sequence of human chorionic somatomammotropin."

RL Arch. Biochem. Biophys. 155:95-112(1973).

RN [9]

RP SEQUENCE OF 27-117.

RX MEDLINE=72016113; PubMed=5286363;

RA Sherwood L.M., Handwerger S., McLaurin W.D., Lanner M.; "Amino-acid sequence of human placental lactogen." Nature New Biol. 233:59-61(1971).

RL [10]

RN ERRATUM.

RA Sherwood L.M., Handwerger S., McLaurin W.D., Lanner M.; Nature New Biol. 235:64-64(1972).

RN [11]

RP INTERCHAIN DISULFIDE BONDS.

RX MEDLINE=79173081; PubMed=438159;

RA Schneider A.B., Kowalski A., Russell J., Sherwood L.M.; "Identification of the interchain disulfide bonds of dimeric human placental lactogen." J. Biol. Chem. 254:3782-3787(1979).

RL [12]

RP FUNCTION: SIMILAR TO THAT OF SOMATOTROPIN.

RX [13]

RP SUBCELLULAR LOCATION: Secreted.

RX [14]

RP MISCELLANEOUS: THE SEQUENCE OF CSH1 IS SHOWN.

RL [15]

RP SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.

CC [16]

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).

CC [17]

CC EMBL: V00573; CAA23836.1; -

CC EMBL: J06289; AAA98747.1; -

CC EMBL: K02401; AAA52115.1; -

CC EMBL: M15694; AAA52116.1; -

CC EMBL: J03071; AAA52551.1; -

CC EMBL: J06118; AAA95621.1; -

CC EMBL: BC002717; AA822717.1; -

CC EMBL: BC005921; AA835921.1; -

CC EMBL: BC020756; AA820756.1; -

CC PIR: A26449; A26449.

CC PIR: C32435; LCHUC.

CC HSP: P01241; I822.

CC Genew: HGNC:2440; CSH1.

CC Genew: HGNC:2441; CSH2.

CC MIN: 150200; -

CC GO: GO:0007365; P:pregnancy; TAS.

CC InterPro: IPR001400; Somatotropin.

CC Pfam: PF00103; hormone; 1

CC PROSITE: PS00256; SOMATOTROPIN_1; 1.

CC PROSITE: PS00338; SOMATOTROPIN_2; 1.

CC Hormone: Placenta; Multigene family; Signal.

CC SIGNAL 1 26

CC CHAIN 27 217 LACTOGEN.

CC DISULFID 79 191

CC DISULFID 208 215

CC DISULFID 208 208

CC DISULFID 215 215

CC VARIANT 3 3

CC VARIANT 104 105

CC CONFLICT 84 84 I -> T (IN REF. 9).

CC CONFLICT 95 95 MISSING (IN REF. 9).

CC CONFLICT 126 116 MISSING (IN REF. 9).

CC CONFLICT 134 136 SDD -> BBS (IN REF. 9).

CC SEQUENCE 217 AA; 25620 MW; 235B0DC7A713F431 CRC64;

CC [18]

CC Query Match 81.18; Score 381; DB 1; Length 217;

CC Best local similarity 62.04; Pred. No. 1.4e-42;

CC Matches 73; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

CC [19]

CC 4 TIPSRLFDNMLRAHRLHOLAIDTQVEFEAVIPKQKYSFLQNPOITS-SFSESPTPS 63

CC 1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

CC 29 TVPSRLFDHMLQAHRAHOLAIDTQVEFEVTPKQKYSFLQDSQTSFCFSDSPTPS 88

CC 64 NREFTQOKSNLELLR-SLLLIQSWLEPVQ 92


```

RL J. Biol. Chem. 260:9574-9579(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN: F21D0; TISSUE= Liver;
RX MEDLINE=9519489; PubMed=8347448;
RA Das P., Meyer L., Seyfert J.N., Brockmann G., Schwert M.
RT "Structure of the growth hormone-encoding gene and its promoter in
   mice."
RL Gene 169:209-213(1996).
CC -1- FUNCTION: Plays an important role in growth control. Its major
CC role in stimulating body growth is to stimulate the liver and
CC other tissues to secrete IGF-1. It stimulates both the
CC differentiation and proliferation of myoblasts. It also stimulates
CC amino acid uptake and protein synthesis in muscle and other
CC tissues.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
DR EMBL: X02991; CAA26650.1;
DR FMBL: 245653; CAA86058.1;
DR PIR: B23911; STMS.
DR HSSP: P01246; IBS7.
DR MGD: MGI:95707; Ch.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone; 1.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW Hormone; Pituitary; Signal.
FT SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 27 216 SOMATOTROPIN.
FT DISULFID 78 189 BY SIMILARITY.
FT DISULFID 206 214 BY SIMILARITY.
SQ SEQUENCE 216 AA; 24716 MW; 9866FAJAE25D65FC CRC64;
Query Match 64.8%; Score 304.5; DB 1; Length 216;
Best Local Similarity 64.8%; Pred. No. 2e-24;
Matches 59; Conservative 14; Mismatches 17; Indels 1; Gaps 1;
QY 2 FPTPLSRFDNAMLRAHRLAQAFDYQFEERAYIPKEOKYSLFLONPQTSLSFSISPT 61
DB 27 FPAVPLSLSFANVLRADHQLAADTYKFEERAYIPKEOKYSLFLONPQTSLSFSISPT 65
QY 62 PSNREETQOKSNLELLRISLLIQSWLEPVQ 92
DB 86 PTGKEAQQRDMELLRFSLLIQSWLGPVQ 116
RESULT 13
SOMA_HORSE
ID SOMA_HORSE STANDARD; PRT: 216 AA.
AC P01245;
DI 21-JUL-1986 (rel. 31, Created)
DI 01-NOV-1995 (rel. 32, Last sequence update)
DI 28-FEB-2003 (rel. 41, Last annotation update)
DE Somatotropin precursor (Growth hormone).
GN GH1.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutrachia; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE= Pituitary;
RX MEDLINE=9426171; PubMed=8206352;
RA Ascacio-Martinez J.A., Barreira-Saldana H.A.;

```

```

"Sequence of a cDNA encoding horse growth hormone.";
Gene 143:299-300(1994).
RN [2]
RP SEQUENCE OF 27-216.
RX MEDLINE=77005410; PubMed=965151;
RA Zakin M.M., Poskus E., Langton A.A., Ferrara P., Santone J.A.,
RA Dellacha J.M., Paladini A.C.;
PT "Primary structure of equine growth hormone.";
RL Ist. C. Pept. Protein Res. 8:435-444(1976).
RN [3]
RP PRELIMINARY SEQUENCE OF 27-216.
RX MEDLINE=74020362; PubMed=4747843;
RA Zakin M.M., Poskus E., Dellacha J.M., Paladini A.C., Santone J.A.;
RT "The amino acid sequence of equine growth hormone.";
RL FEBS Lett. 34:353-355(1973).
RN [4]
RP SEQUENCE OF 68-95 AND 183-216.
RA Zakin M.M., Poskus E., Dellacha J.M., Paladini A.C., Santone J.A.;
RT "Amino acid sequences around the cysteine residues in equine growth
   hormone.";
RL FEBS Lett. 25:77-82(1972).
RN [5]
RP SEQUENCE OF 202-216.
RX MEDLINE=68368390; PubMed=4876100;
RA Gliver L., Hartree A.S.;
RT "Amino acid sequences around the cysteine residues in horse growth
   hormone.";
RL Biochem. J. 109:19-24(1968).
CC -1- FUNCTION: Plays an important role in growth control. Its major
CC role in stimulating body growth is to stimulate the liver and
CC other tissues to secrete IGF-1. It stimulates both the
CC differentiation and proliferation of myoblasts. It also stimulates
CC amino acid uptake and protein synthesis in muscle and other
CC tissues.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
EMBL: X02929; AAA21027.1;
DR HSSP: P01246; IBS7.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone; 1.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW Hormone; Pituitary; Signal.
FT SIGNAL 1 26
FT CHAIN 27 216 SOMATOTROPIN.
FT DISULFID 78 189
FT DISULFID 206 214
SQ SEQUENCE 216 AA; 24423 MW; 37AB3173834D11AC CRC64;
Query Match 64.4%; Score 302.5; DB 1; Length 216;
Best Local Similarity 64.8%; Pred. No. 2e-24;
Matches 59; Conservative 14; Mismatches 17; Indels 1; Gaps 1;
QY 2 FPTPLSRFDNAMLRAHRLAQAFDYQFEERAYIPKEOKYSLFLONPQTSLSFSISPT 61
DB 27 FPAVPLSLSFANVLRADHQLAADTYKFEERAYIPKEOKYSLFLONPQTSLSFSISPT 65
QY 62 PSNREETQOKSNLELLRISLLIQSWLEPVQ 92
DB 86 PTGKEAQQRDMELLRFSLLIQSWLGPVQ 116
RESULT 14
SOMA_RABIT

```

ID	SOMA_RAB1T	STANDARD:	SWT:	216 AA.
AC	P46377			
DT	01-NOV-1995	(Rel. 32, Created)		
DT	01-NOV-1995	(Rel. 32, Last sequence update)		
DT	28-FEB-2003	(Rel. 41, Last annotation update)		
DE	Somatotropin precursor (Growth hormone)			
GN	GHI			
CS	Oryctolagus cuniculus (Rabbit)			
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
CC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus			
CC	NCBI_TaxID=9986;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=New Zealand white;			
RC	MEDLINE=96011643; PubMed=7590276;			
RA	Wallis O.C.; Wallis M.;			
RT	"Cloning and characterisation of the rabbit growth hormone-encoding gene."			
RL	Gene 163:253-256(1995).			
CC	!- FUNCTION: Plays an important role in growth control. Its major role in stimulating body growth is to stimulate the liver and other tissues to secrete IGF-1. It stimulates both the differentiation and proliferation of myoblasts. It also stimulates amino acid uptake and protein synthesis in muscle and other tissues.			
CC	!- SUBCELLULAR LOCATION: Secreted.			
CC	!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL: Z36127; CAA86287.1; -			
DR	PIR: S49483; S49483.			
DR	HSSP: P01246; 1BSC.			
DR	InterPro: IPR001400; Somatotropin.			
DR	Pfam: PF00103; hormone; 1.			
DR	PRINTS: PR00835; SOMATOTROPIN.			
DR	PROSITE: PS00266; SOMATOTROPIN_1; 1.			
DR	PROSITE: PS00338; SOMATOTROPIN_2; 1.			
KW	Hormone; Pituitary; Signal.			
FT	SIGNAL 1 26 POTENTIAL.			
FT	CHAIN 27 216 SOMATOTROPIN.			
FT	DISULFID 78 199 BY SIMILARITY.			
FT	DISULFID 205 214 BY SIMILARITY.			
SEQ	SEQUENCE 216 AA: 24433 Md; 6EC19745199F9575 CRC64;			
Query Match	64.4%; Score 302.5; DR 1; Length 216;			
Best Local Similarity	64.8%; Pred. NC. 2e-24;			
Matches	59; Conservative 14; Mismatches 17; Indels 1; Gaps 1;			
QY	2 EPI:PIRSF:FDNAMI;PABSLHQLAPTYQGEFAYIPKCKYKSYFLNPGQSLSESESLD 61			
DD	27 FPMR:SSUFANAVLRQAQLHQIAADTYKEFAYIFEGQYIS TQMAAGNCTSETTGA 43			
QY	62 FPMR:ETQOKSNLELLRISILLICSMIRPVA 92			
DD	86 PTKCKDAQOBSMELLRFSLILQCSMLGIVG 116			
RESULT 15				
SOMA_RAT				
ID	SOMA_RAT	STANDARD:	PRT:	216 AA.
AC	P01244;			
DT	21-JUL-1986	(Rel. 01, Created)		
DT	21-JUL-1986	(Rel. 01, Last sequence update)		
DT	28-FEB-2003	(Rel. 41, Last annotation update)		
DE	Somatotropin precursor (Growth hormone)			
GN	GHI OR GH.			

QY 2 FETIFLSRLFONAMLRARERHOLAFTDYQSEFEAYIPKQKYSFLONFQTSISESESIPI 61
Db 27 PPAMPSSSLFANAVLRAOHLHOLAATYKSEFERAYIPKQKYS TONACAAKSESESIPIA 62
QY 62 PSNRETCOKSNEELRISLLIUSWLEPVQ 92
Db 86 PTGKEEAQORTONELLRFESALLIUSWIGPVQ 116

Search completed: September 15, 2003, 12:01:29
Job time : 8.57348 secs

CC A and B chains separated by a 34 residue peptide sequence. This insulin
 CC precursor can be a component of chimeric proteins which additionally
 CC contains an N-terminal fragment of human growth hormone (hGH); and a
 CC cleavable peptide linker (AA42857). The hGH portion of the chimeric
 CC protein acts as an intramolecular chaperone (IMC) for the insulin
 CC precursor, enabling it to fold correctly. The cleavable peptide linker
 CC has a C-terminal Arg residue which enables the hGH portion of the
 CC chimeric protein to be removed after folding has taken place. Production
 CC of recombinant human insulin via an hGH-proinsulin chimeric protein can
 CC provide human insulin with correctly linked cysteine bridges with
 CC fewer necessary procedural steps, and hence resulting in a higher yield
 CC of human insulin. The IMC sequences not only protect insulin sequences
 CC from intracellular degradation by a microorganism host, but also promote
 CC the folding of the fused insulin precursor, facilitate the solubility of
 CC the fusion protein and decrease the intermolecular interactions among
 CC the fusion proteins, thus allowing folding of the fused insulin precursor
 CC at commercially useful high concentrations. The procedural steps of
 CC cyanogen bromide cleavage, oxidative sulphatolysis and related
 CC purification steps can thus be eliminated, along with the use of high
 CC concentrations of mercaptan or the use of hydrophobic absorbent resins.
 XX
 SQ Sequence 86 AA:

Query Match 100.0%; Score 463; DP 20; Length 86;

Best Local Similarity 100.0%; Pred. No. 8,4e-43;

Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVAEALVYVCGRGFFFTPKTPREAEDELQVGLGSGPGAGS;QPLALRG 60
 DB 1 FVNHLCGSHLVAEALVYVCGRGFFFTPKTPREAEDELQVGLGSGPGAGS;QPLALRG 60

QY 61 SLQKRGIVEQCCTSCSLYLENYCN 86

DB 61 SLQKRGIVEQCCTSCSLYLENYCN 86

RESULT 4

AA012770

ID AAB12770 standard; protein: 85 AA.

XX AAB12770:

XX AAB12770:

DT 22-NOV-2000 (first entry)

XX Human proinsulin protein sequence SEQ ID NO:2.

XX Human; insulin-like growth factor 1; IGF-1; proinsulin; insulin.
 KW mutant; variant; insulin-like growth factor binding protein; IGF-1;
 KW IGF-3; antidiabetic; neuroprotective; anorectic; tranquiliser;
 KW vulnery; anorectic; cardiant; nephrotropic; dermatological; antiHIV;
 KW antiviral; hyperglycaemia; obesity; lung disease; glomerulonephritis;
 KW interstitial nephritis; Turner's syndrome; Laron's syndrome;
 KW short stature; increased fat mass-to-lean ratio; immunological disorder;
 KW peripheral neuropathy; multiple sclerosis; muscular dystrophy;
 KW catabolic state; trauma; wounding; infection; HIV; skin disorder;
 KW human immunodeficiency virus; diabetes; heart dysfunction;
 KW kidney disorder; whole body growth disorder.

XX Homo sapiens.

XX WC2000040612-A1.

XX 13-JUL-2000.

XX 05-JAN-2000; 2000WO-US00151.

XX 06-JAN-1999; 99US-0115010.

XX (GETH) GENENTECH INC.

XX Dubaquié Y, Lowman H;

XX WPI: 2000-455955/40.

XX Novel insulin-like growth factor (IGF) 1 mutants that selectively bind
 PT to IGF binding protein (IGFBP)-1 or IGFBP-3, used to improve the
 PT half-lives of IGF-1 and insulin -

XX Disclosure: Page 44; 48pp; English.

XX The present invention describes an insulin-like growth factor (IGF)-1
 CC variant (I), where an amino acid at position 3, 4, 5, 7, 10, 14, 17, 23,
 CC 24, 25, 43, 49 or 53, optionally in combination with an amino acid at
 CC position 12 and/or 16 of the native human IGF-1 sequence, is replaced
 CC with an alanine, glycine, or a serine residue. The residue at position 7
 CC may be replaced by any amino acid. (I) can have antidiabetic, cardiac,
 CC neuroprotective, anorectic, tranquiliser, vulnery, anorectic,
 CC nephrotropic, dermatological, antiHIV and antiviral activities. The IGF-1
 CC mutants are used in any methods where IGFs or insulin are used, e.g. in
 CC treating hyperglycaemia, obesity-related, neurological, cardiac, renal,
 CC immunological, and anabolic disorders. These disorders include lung
 CC diseases, glomerulonephritis, interstitial nephritis, Turner's syndrome,
 CC Laron's syndrome, short stature, increased fat mass-to-lean ratio,
 CC immunological disorders, peripheral neuropathy, multiple sclerosis,
 CC muscular dystrophy, catabolic states, trauma, wounding, infection,
 CC human immunodeficiency virus (HIV), wounds, skin disorders, diabetes,
 CC heart dysfunctions, kidney disorders, and whole body growth disorders.
 CC They can also be used for increasing serum and tissue levels of
 CC biological active IGF or insulin in a mammal. The IGF-1 mutants improve the
 CC half-lives of IGF-1 and insulin. The present sequence represents the
 CC native human proinsulin protein sequence, which is given in the
 CC exemplification of the present invention.

XX Sequence 86 AA:

Query Match 100.0%; Score 463; DP 20; Length 86;

Best Local Similarity 100.0%; Pred. No. 8,4e-43;

Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVAEALVYVCGRGFFFTPKTPREAEDELQVGLGSGPGAGS;QPLALRG 60
 DB 1 FVNHLCGSHLVAEALVYVCGRGFFFTPKTPREAEDELQVGLGSGPGAGS;QPLALRG 60

QY 61 SLQKRGIVEQCCTSCSLYLENYCN 86

DB 61 SLQKRGIVEQCCTSCSLYLENYCN 86

RESULT 5

AA048215

ID AAM48215 standard; protein: 86 AA.

XX AAM48215:

XX AAM48215:

DT 16-MAR-2002 (first entry)

XX Human proinsulin.

XX Antirheumatic; antiarthritic; osteopathic; cartilage disorder;
 KW insulin-like growth factor; IGF; binding protein; IGFBP;
 KW rheumatoid arthritis; osteoarthritis; proinsulin; human.

XX Homo sapiens.

XX WC200187323-A2.

XX 22-NOV-2001.

XX 16-MAY-2001; 2001WO-US15934.

XX 16-MAY-2000; 2000US-204490F.

XX 15-NOV-2000; 2000US-248985P.

XX (GETH) GENENTECH INC.

XX Dubaquié Y, Filvarolf EH, Lowman HB;

```

XX DR WPI: 2002-082942/11.
XX PT Treating cartilage disorders including cartilage damage by delivery of
XX PT degenerative cartilaginous disorders, by contacting cartilage with
XX PT insulin-like growth factor analog with altered affinity for TGF-beta1
XX PT proteins.
XX PS Disclosure: Fig 16: 136pp; English.
XX CC The present invention relates to a method for treating cartilage
XX CC disorders. The method comprises contacting cartilage with an active agent
XX CC such as insulin like growth factor (IGF-1) analog with a binding affinity
XX CC preference for IGF binding protein-3 (IGFBP-3) over IGF-1 than an IGF-1
XX CC analog with a binding affinity preference for IGFBP-3 over IGFBP-1, or a
XX CC IGFBP displacer peptide that prevents the interaction of IGF with an
XX CC IGFBP and does not bind to human IGF receptor. The method is useful for
XX CC treating cartilage disorders (CD), including degenerative CD, articular
XX CC CD such as rheumatoid arthritis and osteoarthritis. The present sequence
XX CC is human proinsulin, which was used to illustrate the invention.
XX SQ Sequence 85 AA:

Query Match 100.0%; Score 463; E: 23; Length 86;
Best Local Similarity 100.0%; Pred. No. 8.4e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FVNQLCGSHLVEALYVCGERGFFYTPKTRREAEDLVQGVGGGPGAGSGLOPLALEG 60
DB 1 FVNQLCGSHLVEALYVCGERGFFYTPKTRREAEDLVQGVGGGPGAGSGLOPLALEG 60
OY 61 SLQKRGIVEQCCTSIQSIYOLENYCN 86
DB 61 SLQKRGIVEQCCTSIQSIYOLENYCN 86

RESULT 6
AAP20036
ID AAP20036 standard; Protein: 87 AA.
AC AAP20036;
XX 25-MAR-2003 (updated)
DI 22-JUL-1992 (first entry)
XX Human proinsulin.
XX Proinsulin.
XX Homo sapiens.
XX EP55942-A.
XX 14-JUL-1982.
XX 31-DEC-1982; 81EP-030619C.
XX 02-JAN-1981; 81US-022201C.
XX 23-JUL-1981; 81US-028607C.
XX 02-JAN-1982; 82US-022201C.
XX 03-MAR-1982; 82US 0354287.
XX (CNYN-) STATE UNIV NEW YORK.
XX Inouye M, Nakamura K;
XX WPI: 1982-59775E/29 (59775E).
XX N-PSDB; AAN20041.
XX Plasmid cloning vehicles - useful for transforming bacterial
XX hosts to produce eukaryotic polypeptide(s).
XX Disclosure: Fig 27: 114pp; English.

```

```

XX CC The sequence comprises human proinsulin.
XX CC (Updated on 25-MAR-2003 to correct PR field.)
XX SQ Sequence 87 AA:

Query Match 100.0%; Score 463; E: 3; Length 87;
Best Local Similarity 100.0%; Pred. No. 8.5e-43;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FVNQLCGSHLVEALYVCGERGFFYTPKTRREAEDLVQGVGGGPGAGSGLOPLALEG 60
DB 2 FVNQLCGSHLVEALYVCGERGFFYTPKTRREAEDLVQGVGGGPGAGSGLOPLALEG 61
OY 61 SLQKRGIVEQCCTSIQSIYOLENYCN 86
DB 62 SLQKRGIVEQCCTSIQSIYOLENYCN 87

RESULT 7
AAP40217
ID AAP40217 standard; Protein: 87 AA.
AC AAP40217;
XX 25-MAR-2003 (updated)
DI 12-FEB-1992 (first entry)
XX Sequence of the 32 N-terminal AAs of proinsulin.
XX Hormone; cloning vector; phage resistant.
XX Homo sapiens.
XX Key Location/Qualifiers
XX Region 2..31
XX Region /label= B-chain
XX Region 32..66
XX Region /label= C-chain
XX Region 67..87
XX Region /label= A-chain
XX GB2126237-A.
XX 21-MAR-1984.
XX 01-SEP-1983; 83GS-0023468.
XX 03-SEP-1982; 82US-041429C.
XX 05-SEP-1984; 84US-0647338.
XX (ECL) LILLY & CO ELI.
XX Hershberger CL, Rostock PR;
XX WPI: 1984-070793/12.
XX N-PSDB; AAN40179.
XX Protecting bacteria from phage infection - by transformation with
XX cloning vector contg. segment with restriction and modification
XX activity
XX Example: Fig 10: 28pp; English.
XX Plasmid pTh alpha 1 was constructed by inserting a synthesised gene
XX for thymosin alpha 1 (AAN40178) into plasmid pBR322. It is used for
XX the construction of pPrp24. The inventors claim a method for
XX protecting bacteria from phage infection - by transformation with
XX cloning vector contg. segment with restriction and modification
XX activity. Prodn. of plasmid pPr 26 or pPr27 which uses pPrp24; and
XX prodn. of plasmid pPr29 which uses a synthetic gene coding for the
XX 32 N-terminal AAs of proinsulin (see AAN40179).
XX (Updated on 25-MAR-2003 to correct PA field.)

```

```

XX SQ Sequence 87 AA;
    Query Match 100.0%; Score 463; DB 6; Length 87;
    Best Local Similarity 100.0%; Pred. No. 8.5e-43;
    Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FVNHGCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQLGGGPGAGSLQPLALEG 60
    IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
DB 2 FVNHGCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQLGGGPGAGSLQPLALEG 61
    IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
QY 61 SLOKRGIVEQCCTSGCSLYOLENYCN 86
    IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
DB 62 SLOKRGIVEQCCTSGCSLYOLENYCN 87
    IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII

RESULT 8
AAP50127
ID AAP50127 standard; Protein: 67 AA.
XX
AC AAP50127;
XX
DT 25-MAR-2003 (updated)
DT 16-AUG-2002 (updated)
DT 30-SEP-1991 (first entry)
XX
DE Sequence of the 32 N-terminal AAs of proinsulin.
XX
KW Selectable vector; autonomously replicating vector;
    expression vector.
XX
OS Homo sapiens.
XX
QS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 2..31 /label= A chain.
FT Region 32..66 /label= B chain.
FT Region 67..87 /label= C chain.
FT Region /label= A chain.
XX
PN EP154539-A.
XX
PD 11-SEP-1985.
XX
PF 04-MAR-1985; 85EP-0301469.
XX
PR 06-MAR-1984; 84JUS-0586592.
XX
PR 01-APR-1985; 85CN-0107555.
XX
PA (EJLJ ) LILLY & CO ELI.
XX
PI Schoner R, Schoner B;
XX
DR WPI; 1985-224921/37.
DR N-PSDB; AAN50152.
XX
PT New recombinant DNA expression vector with autonomous
    replication and on transcription generating polycistronic mRNA
XX
PS Example; Fig 14; 118pp; English.
XX
CC The inventors claim a process for preparing selectable and
    autonomously replicating recombinant DNA expression vectors which
    comprise 1) a transcriptional and translational activating sequence
    which is in the reading frame of a nucleotide sequence which codes
    for a peptide or polypeptide; 2) a translational stop signal; 3) a
    translational start signal which is in the reading frame of a
    nucleotide sequence that codes for a functional polypeptide;
    and 4) an additional translational stop signal. The peptide or
    polypeptide coding sequence codes for 2-20 AAs, esp. AAP50122-p50125.
    The functional polypeptide is esp. growth hormone, human insulin.

```

```

CC Interferon and human tissue plasminogen activator.
CC (Updated on 16-AUG-2002 to add missing OS field.)
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 87 AA;
    Query Match 100.0%; Score 463; DB 6; Length 87;
    Best Local Similarity 100.0%; Pred. No. 8.5e-43;
    Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FVNHGCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQLGGGPGAGSLQPLALEG 60
    IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
DB 2 FVNHGCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQLGGGPGAGSLQPLALEG 61
    IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
QY 61 SLOKRGIVEQCCTSGCSLYOLENYCN 86
    IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
DB 62 SLOKRGIVEQCCTSGCSLYOLENYCN 87
    IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII

RESULT 9
AAP50060
ID AAP50060 standard; Protein: 87 AA.
XX
AC AAP50060;
XX
DT 25-MAR-2003 (updated)
DT 16-AUG-2002 (updated)
DT 11-NOV-1991 (first entry)
XX
DE Synthetic proinsulin.
XX
KW Proinsulin; vector; proteinaceous granule;
    Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 1..30 /label= B chain.
FT Region 31..65 /label= C chain.
FT Region 66..86 /label= A chain.
FT Region /label= A chain.
XX
PN EP159123-A.
XX
PD 23-OCT-1985.
XX
PF 04-MAR-1985; 85EP-0301469.
XX
PR 06-MAR-1984; 84JUS-0586592.
PR 26-JUL-1984; 84JUS-0634920.
PR 31-JAN-1985; 85JUS-0697090.
XX
PA (EJLJ ) LILLY & CO ELI.
XX
PI Hsiung HM, Schiner RG, Schoner BE;
XX
DR WPI; 1985-265030/43.
DR N-PSDB; AAN50082.
XX
PT New selectable and autonomously replicating DNA expression vector -
    useful in producing proteinaceous granules in cell transformants, esp.
XX
PS Disclosure; Fig. 14; 115pp; English.
XX
CC The synthetic proinsulin gene is expressed in a new selectable and
    autonomously replicating recombinant DNA expression vector
    comprising a runaway replicon and a transcriptional and
    translational activating sequence in the reading frame of the
    proinsulin coding sequence, the sequence contg. a translational
    stop signal. Host cells contg. the vector, which is esp. p.asmid
    pz1103, are cultured, and proinsulin is produced as a highly
    homogeneous species of proteinaceous granule. The granule can be

```

CC readily isolated from cell lysates and is stable on washing with
 CC urea or detergent soaps, at low osmotic. The analogue contains at
 CC least 50% of proinsulin and all isolation operations are
 CC simplified.
 CC (Updated on 16-AUG-2002 to add missing GS field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 87 AA;

Query Match 100.0%; Score 463; DB 6; Length 87;
 Best Local Similarity 100.0%; Pred. No. 8.5e-43;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 61
 DB 2 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 61
 QY 61 SLOKRGIVEQCTSIQSLYLENYCN 86
 DB 62 SLOKRGIVEQCTSIQSLYLENYCN 87

RESULT 10
 AAP61090 standard; Protein: 87 AA.
 XX
 AC AAP61090;
 XX
 DC 28-FEB-1992 (first entry)
 XX
 DE Sequence encoded by the structural gene for human proinsulin.
 XX
 KW Recombinant plasmid; E.coli expression vector; secretion vector.
 XX
 OS Homo sapiens.
 XX
 PN US4624926-A.
 XX
 PD 25-NOV-1985.
 XX
 PF 03-MAR-1982; 82US-0354287.
 XX
 PR 03-MAR-1982; 82US-0354287.
 PR 02-JAN-1981; 81US-0222010.
 PR 23-JUL-1981; 81US-0286076.
 XX
 XX (UNY-) UNIV OF NEW YORK.
 XX
 PI Inouye M, Nakamura K;
 XX
 DR WPI: 1986-331802/50.
 DR N-PSDB; AAN60872.
 XX
 PR New recombinant plasmid(s) - conts. DNA sequences encoding
 PR exogenous polypeptide and outer membrane protein of E.coli;
 XX
 PS Example: Fig 27; 44pp; English.

XX The inventors claim new recombinant plasmids contg. a DNA sequence
 CC encoding a polypeptide, which is foreign to E.coli, in reading phase
 CC with a DNA SQ. coding for at least one functional fragment derived
 CC from an outer membrane lipoprotein gene of E.coli. The foreign gene
 CC may be for human insulin. The lipoprotein gene functional fragment
 CC may be the promoter, the 5'-UTR, the 3'-UTR or the transcription
 CC termination signal provided that it includes at least the promoter.
 XX
 SQ Sequence 87 AA;
 Query Match 100.0%; Score 463; DB 6; Length 87;
 Best Local Similarity 100.0%; Pred. No. 8.5e-43;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 61

DB 2 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 61
 QY 61 SLOKRGIVEQCTSIQSLYLENYCN 86
 DB 62 SLOKRGIVEQCTSIQSLYLENYCN 87
 RESULT 11
 AAR32367
 IC AAR32367 standard; Protein: 87 AA.
 XX
 AC AAR32367;
 XX
 DT 25-MAR-2003 (updated)
 DT 16-JUN-1993 (first entry)
 XX
 DE Proinsulin protein sequence.
 XX
 KW Human; proinsulin; vector; pUC19; pPINS; CAT; pUC-CAT-proinsulin;
 KW insulin analogue; type I; type II; diabetes.
 XX
 OS Synthetic.

XX
 PN WO9303174-A1.
 XX
 PD 18-FEB-1993.
 XX
 PF 31-JUL-1992; 92WO-0506451.
 XX
 PR 08-AUG-1991; 91US-0741938.
 PR 30-JUL-1992; 92US-0918953.
 XX
 PA (P12) PEIZER INC.
 PA (SCIO) SCIOS INC.
 PI Andy RJ, Larson EK;
 XX
 DR WPI: 1993-076530/09.
 DR N-PSDB; AAO37003.
 XX
 PR New hepato selective and peripheral selective human insulin
 PR analogues - and their corresp. DNA, for treatment of type I and
 PR type II diabetes
 XX
 PS Disclosure: Fig 2b; 54pp; English.

XX This sequence represents human proinsulin and was decoded from
 CC the sequences given in AAO36996-7001. The cDNA fragment coding for
 CC proinsulin was inserted into plasmid vector pUC19 and digested with
 CC KpnI and HindIII. This resulted in the formation of the vector pPINS.
 CC A fragment encoding amino acids 1-73 of CAT (see AAO37002) was inserted
 CC into pPINS to give a plasmid which contained DNA sequences which coded
 CC for amino acids 1-73 of CAT, an 8 amino acid linker sequence and human
 CC proinsulin. This plasmid, pUC-CAT-proinsulin, could be used in the
 CC formation of insulin analogues which may be used in the treatment of
 CC types I and II diabetes.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 87 AA;
 Query Match 100.0%; Score 463; DB 14; Length 87;
 Best Local Similarity 100.0%; Pred. No. 8.5e-43;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
 DB 2 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 61
 QY 61 SLOKRGIVEQCTSIQSLYLENYCN 86
 DB 62 SLOKRGIVEQCTSIQSLYLENYCN 87

RESULT 12

AA07682
ID AAR07682 standard; protein: 86 AA.

XX AC AAR07682;
XX
XX 25-MAR-2003 (updated)
DT 09-JAN-2003 (updated)
DT 13-FEB-1991 (first entry)
XX DE Modified human insulin precursor.
XX DE Human insulin precursor; cathepsin C.
XX KW Homo sapiens.
OS
XX Key Location/Qualifiers
FH Peptide 1..12
FT Peptide /label-N-terminal initiating dipeptide
FT Peptide 1..32
FT Peptide /label-native human insulin B-chain
FT Peptide 33..67
FT Peptide /label-natural connecting peptide of human
FT Peptide 68..88
FT Peptide /label-native human insulin A-chain
XX EP397420-A.
XX 14-NOV-1990.
XX 04-MAY-1990; 80EP-0304890.
XX 09-MAY-1989; 89US-0349471.
XX (ELIL) LILLY & CO ELI.
XX
XX Becker GW, Furman TC, Mackellar WT, McKenough JR;
XX WPI: 1990-343372/45.
XX Human insulin precursor - contg. Met-Tyr or Met-Tyr initiating
XX dipeptide for controlled removal by cathepsin C
XX
XX Disclosure; page 3; 8pp; English.
XX
XX This modified human insulin precursor comprises an N-terminal
XX initiating dipeptide, chosen from Met-Tyr or Met-Arg, which does
XX not define a cathepsin C dipeptide removal stop point. This
XX dipeptide is linked to the natural human insulin B-chain, natural
XX human proinsulin connecting peptide and natural human insulin A-
XX chain. Dipeptide removal is carefully controlled to obtain the
XX desired prod. without further degradation occurring, irrespective
XX of whether the next dipeptide in the sequence defines a cathepsin
XX C stop point.
XX (Updated on 29-JAN-2003 to add miss.-sg OS field.)
XX (Updated on 25-MAR-2003 to correct PA field.)
XX

XX Sequence 86 AA:

Query Match 100.0%; Score 463; DB 1; Length 86;
Best Local Similarity 100.0%; Pred. No. 8.6e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVEALYVGCGERFFYTPKTRAEADLVQGVELGGPGAGSLQPLALEG 69
DB 3 FVNHLCGSHLVEALYVGCGERFFYTPKTRAEADLVQGVELGGPGAGSLQPLALEG 62

QY 61 SLQKRGIVECCCTSGISLYQLENYCN 86

DB 63 SLQKRGIVECCCTSGISLYQLENYCN 88

RESULT 13

AA03855
ID AAR03855 standard; Protein: 88 AA.

XX AC AAR03855;
XX
XX 25-MAR-2003 (updated)
DT 19-JUL-1993 (first entry)
XX DE hpl.
XX
XX Proinsulin; hpl; native; pc2R126S; expression vector; E. coli; human;
XX expression; immunological effect.
XX
XX Homo sapiens.
XX EP534705-A2.
XX 31-MAR-1993.
XX 22-SEP-1992; 92EP-0308601.
XX 24-SEP-1991; 91US-0764655.
XX (ELIL) LILLY & CO ELI.
XX
XX Belagade RM;
XX
XX WPI: 1993-102806/13.
XX N-PSDB; AAQ38310.
XX
XX Expression of low molecular wt. polypeptide(s) e.g. insulin
XX growth factor I - by expressing as deriv. with N-terminal
XX amino acid to provide increased expression levels
XX
XX Disclosure; Page 21-22; 40pp; English.
XX
XX This sequence represents an analogue of native human proinsulin (hpl).
XX The DNA encoding this sequence was used in the construction of the
XX expression vector of the invention. The coding region of the hpl gene
XX was synthesised and was cloned into the expression plasmid pc2R126S
XX (see also AAQ38307). Expression of this gene lead to the inclusion of
XX an extra amino acid (Arg) in the second position from the N-terminal
XX of mature hpl. The extra amino acid provides increased expression
XX levels of the protein and is then cleaved off to avoid undesirable
XX immunological effects when used in humans.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX

XX Sequence 88 AA;

Query Match 100.0%; Score 463; DB 14; Length 88;
Best Local Similarity 100.0%; Pred. No. 8.6e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVEALYVGCGERFFYTPKTRAEADLVQGVELGGPGAGSLQPLALEG 60
DB 3 FVNHLCGSHLVEALYVGCGERFFYTPKTRAEADLVQGVELGGPGAGSLQPLALEG 62

QY 61 SLQKRGIVECCCTSGISLYQLENYCN 86

DB 63 SLQKRGIVECCCTSGISLYQLENYCN 88

RESULT 14

AA020467
ID AAR020467 standard; Protein: 92 AA.

XX AC AAR020467;
XX

XX 25-MAR-2003 (updated)
DT 21-APR-1992 (first entry)
XX


```
Query Match 100.0%; Score 463; DB 4; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.5e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQVELGGGPGAGSLQPLALEG 60
DB 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQVELGGGPGAGSLQPLALEG 60
QY 61 SLOKRGIVEGCGCTSCSLYLENYCN 86
DB 61 SLOKRGIVEGCGCTSCSLYLENYCN 86

RESULT 3
US-09-723-896-2
; Sequence 2, Application US/09723896
; Patent No. 6559443
; GENERAL INFORMATION:
; APPLICANT: Dubaqui, Yves
; APPLICANT: Lowman, Henry
; TITLE OF INVENTION: PROTEIN VARIANTS
; FILE REFERENCE: P1712R
; CURRENT APPLICATION NUMBER: US/09/723,896
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US/99/477,923
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 5
; SEQ ID NO 2
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-723-896-2

Query Match 100.0%; Score 463; DB 4; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.5e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQVELGGGPGAGSLQPLALEG 60
DB 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQVELGGGPGAGSLQPLALEG 60
QY 61 SLOKRGIVEGCGCTSCSLYLENYCN 86
DB 61 SLOKRGIVEGCGCTSCSLYLENYCN 86

RESULT 4
US-09-878-380-1
; Sequence 1, Application US/09878380
; Patent No. 6534281
; GENERAL INFORMATION:
; APPLICANT: FujisBio Inc.
; APPLICANT: KITAJIMA, Sachiko
; APPLICANT: KURANO, Yoshihiro
; APPLICANT: NAKATSUGA, Kaoru
; APPLICANT: NISHIZONO, Isao
; TITLE OF INVENTION: Immunoassay For Measuring Human C-Peptide and its Derivatives
; FILE REFERENCE: 0760-021P
; CURRENT APPLICATION NUMBER: US/09/878,380
; CURRENT FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: JP 2000-174691
; PRIOR FILING DATE: 2000-06-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent in version 3.
; SEQ ID NO 1
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-878-380-1

Query Match 100.0%; Score 463; DB 4; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.5e-47;
```

```
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQVELGGGPGAGSLQPLALEG 60
DB 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQVELGGGPGAGSLQPLALEG 60
QY 61 SLOKRGIVEGCGCTSCSLYLENYCN 86
DB 61 SLOKRGIVEGCGCTSCSLYLENYCN 86

RESULT 5
US-09-134-836-4
; Sequence 4, Application US/09134836
; Patent No. 5956048
; GENERAL INFORMATION:
; APPLICANT: Rubroder, Franz-Josef
; APPLICANT: Keller, Reinhold
; TITLE OF INVENTION: Improved process for obtaining
; TITLE OF INVENTION: insulin precursors having correctly bonded cystine bridge
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farrahaw, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/134,836
; APPLICATION NUMBER: US/09/134,836
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Leslie McDonnell
; REGISTRATION NUMBER: 34,872
; REFERENCE/DOCKET NUMBER: 02481.1600-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 96 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..96
US-09-134-836-4

Query Match 100.0%; Score 463; DB 2; Length 96;
Best Local Similarity 100.0%; Pred. No. 1.6e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQVELGGGPGAGSLQPLALEG 60
DB 11 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQVELGGGPGAGSLQPLALEG 70
QY 61 SLOKRGIVEGCGCTSCSLYLENYCN 86
DB 71 SLOKRGIVEGCGCTSCSLYLENYCN 96

RESULT 6
```

```
US-09-386-303A-4
: Sequence 4, Application US/09/36303A
: Patent No. 6380355
: GENERAL INFORMATION:
: APPLICANT: Rubroder, Franz-Josef
: TITLE OF INVENTION: Improved process for obtaining
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Finnegan, Henderson, Farrelow, Garrett &
: Dunner
: STREET: 1300 7 Street, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20005-3415
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/386,303A
: FILING DATE: 31-Aug-1999
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 09/134,846
: FILING DATE: <Unknown>
: ATTORNEY/AGENT INFORMATION:
: NAME: Leslie McDowell
: REGISTRATION NUMBER: 34,872
: REFERENCE/DOCKET NUMBER: 02481.1609-00000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 408-4500
: TELEFAX: (202) 408-4400
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 96 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: ORIGINAL SOURCE:
: ORGANISM: Escherichia coli
: FEATURES:
: NAME/KEY: Promicin
: LOCATION: 1-96
: SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-386-303A-4

Query Match 100.0%; Score 463; DB 4; Length 96;
Best Local Similarity 100.0%; Pred. No. 1,8e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNCHLGGSHLVYALVVGGRGFFYTPKTRREADLQVGVGLGGPGAGSLQPLALRG 60
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11 FVNCHLGGSHLVYALVVGGRGFFYTPKTRREADLQVGVGLGGPGAGSLQPLALRG 70
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 61 SLOKRGIVEOCCTSCSLQLENYCN 86
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 71 SLOKRGIVEOCCTSCSLQLENYCN 96
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7
US-08-160-376A-4
: Sequence 4, Application US/08/160376A
: Patent No. 5473049
: GENERAL INFORMATION:
: APPLICANT: Obermeyer, Ranier
: APPLICANT: Gerl, Martin
: APPLICANT: Ludwig, Jurgen
: APPLICANT: Sabel, Walter
```

```
US-08-160-376A-4
: TITLE OF INVENTION: Process For Obtaining Proinsulin
: TITLE OF INVENTION: Possessing Correctly Linked
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Kenneth A. Genoni, Esq.
: STREET: Rt. 202-206 No. 5473049th/P.O. Box 2500
: CITY: Somerville
: STATE: New Jersey
: COUNTRY: U.S.A.
: ZIP: 08876-1258
: COMPUTER READABLE FORM:
: MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
: COMPUTER: IBM 386
: OPERATING SYSTEM: WINDOWS 3.1
: SOFTWARE: WORDPERFECT 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/160,376A
: FILING DATE: December 1, 1993
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GE P 4240420.7
: FILING DATE: December 2, 1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Barbara V. Maurer, Esq.
: REGISTRATION NUMBER: 31,287
: REFERENCE/DOCKET NUMBER: HUE 92/P 384
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (908) 231-4079
: TELEFAX: (908) 231-2255
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 97 Amino Acids
: TYPE: Amino Acid (AA)
: TOPOLOGY: not relevant
US-08-160-376A-4

Query Match 100.0%; Score 463; DB 1; Length 97;
Best Local Similarity 100.0%; Pred. No. 1,8e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNCHLGGSHLVYALVVGGRGFFYTPKTRREADLQVGVGLGGPGAGSLQPLALRG 60
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 12 FVNCHLGGSHLVYALVVGGRGFFYTPKTRREADLQVGVGLGGPGAGSLQPLALRG 71
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 61 SLOKRGIVEOCCTSCSLQLENYCN 86
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 72 SLOKRGIVEOCCTSCSLQLENYCN 97
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
US-08-950-720A-11
: Sequence 11, Application US/08950720A
: Patent No. 6046028
: GENERAL INFORMATION:
: APPLICANT: Conklin, Darrell C.
: APPLICANT: Lottor-Day, Catherine E.
: APPLICANT: Lok, Si
: APPLICANT: Jaspers, Stephen R.
: TITLE OF INVENTION: INSULIN HOMOLOG
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: ZymoGenetics, Inc.
: STREET: 1201 Eastlake Avenue East
: CITY: Seattle
: STATE: WA
: COUNTRY: USA
: ZIP: 98102
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
```

1 CURRENT APPLICATION DATA:
2 APPLICATION NUMBER: US/03/950,720A
3 FILING DATE:
4 CLASSIFICATION: 435
5 PRIOR APPLICATION DATA:
6 APPLICATION NUMBER:
7 FILING DATE:
8 ATTORNEY/AGENT INFORMATION:
9 NAME: Sawislak, Deborah A.
10 REGISTRATION NUMBER: 37,638
11 REFERENCE/DOCKET NUMBER: 95-09
12 TELECOMMUNICATION INFORMATION:
13 TELEPHONE: 206-442-6672
14 TELEFAX: 206-442-6678
15
16 INFORMATION FOR SEQ ID NO: 1:
17 SEQUENCE CHARACTERISTICS:
18 LENGTH: 120 amino acids
19 TYPE: amino acid
20 STRANDEDNESS: single
21 TOPOLOGY: linear
22 MOLECULE TYPE: No. 6046028a
23
24 US-08-950-720A-11

Query Match 100.0% Score 463; DB 3; Length 110;
Best Local Similarity 100.0% Pred. No. 2.1e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLYEALVCGSGFFYPKTRREAEADLVGGVCGAGSLOPLALEG 60
DB 25 FVNHLCGSHLYEALVCGSGFFYPKTRREAEADLVGGVCGAGSLOPLALEG 60
QY 61 SLOKRGIVEQCCTSCSLYLENYCN 86
DB 85 SLOKRGIVEQCCTSCSLYLENYCN 110

RESULT 9

US-08-589-028-2
1 Sequence 2, Application US/085:9028
2 Patent No. 6087129
3 GENERAL INFORMATION:
4 APPLICANT: Newgard, Christopher B.
5 APPLICANT: Haiban, Philippe A.
6 APPLICANT: No. 6087129mington, Karl D.
7 APPLICANT: Clark, Samuel A.
8 APPLICANT: Thigpen, Anice E.
9 APPLICANT: Quade, Christian
10 APPLICANT: Kruse, Fred
11 TITLE OF INVENTION: Recombinant Expression of Proteins from
12 TITLE OF INVENTION: Secretary Cell Lines
13 NUMBER OF SEQUENCES: 50
14 CORRESPONDENCE ADDRESS:
15 ADDRESSEE: Arnold, White & Durkee
16 STREET: P. O. Box 4433
17 CITY: Houston
18 STATE: TX
19 COUNTRY: USA
20 ZIP: 77210-4433
21 COMPUTER READABLE FORM:
22 MEDIUM TYPE: Floppy disk
23 COMPUTER: IBM PC compatible
24 OPERATING SYSTEM: PC-DOS/MS-DOS
25 SOFTWARE: PatentIn Release #1.0, Version #1.30
26 CURRENT APPLICATION DATA:
27 APPLICATION NUMBER: US/08/589,028
28 FILING DATE: Concurrently Herewith
29 CLASSIFICATION: 435
30 ATTORNEY/AGENT INFORMATION:
31 NAME: Highlander, Steven L.
32 REGISTRATION NUMBER: 47,642
33 REFERENCE/DOCKET NUMBER: UTSD:426\HYL
34 TELECOMMUNICATION INFORMATION:
35 TELEPHONE: 512/474-7577
36 TELEFAX: 512/474-7577
37 INFORMATION FOR SEQ ID NO: 2:
38 SEQUENCE CHARACTERISTICS:
39 LENGTH: 110 amino acids
40 TYPE: amino acid
41 STRANDEDNESS:

1 TELEPHONE: (512) 418-3000
2 TELEFAX: (512) 474-7577
3 INFORMATION FOR SEQ ID NO: 2:
4 SEQUENCE CHARACTERISTICS:
5 LENGTH: 110 amino acids
6 TYPE: amino acid
7 STRANDEDNESS:
8 TOPOLOGY: linear
9 US-08-589-028-2

Query Match 100.0% Score 464; DB 3; Length 110;
Best Local Similarity 100.0% Pred. No. 2.1e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLYEALVCGSGFFYPKTRREAEADLVGGVCGAGSLOPLALEG 60
DB 25 FVNHLCGSHLYEALVCGSGFFYPKTRREAEADLVGGVCGAGSLOPLALEG 60
QY 61 SLOKRGIVEQCCTSCSLYLENYCN 86
DB 85 SLOKRGIVEQCCTSCSLYLENYCN 110

RESULT 10

US-08-784-582-2
1 Sequence 2, Application US/08784582
2 Patent No. 6110707
3 GENERAL INFORMATION:
4 APPLICANT: Newgard, Christopher B.
5 APPLICANT: Haiban, Philippe A.
6 APPLICANT: No. 6110707mington, Karl D.
7 APPLICANT: Clark, Samuel A.
8 APPLICANT: Thigpen, Anice E.
9 APPLICANT: Quade, Christian
10 APPLICANT: Kruse, Fred
11 APPLICANT: McGarity, Dennis
12 TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
13 TITLE OF INVENTION: SECRETORY CELL LINES
14 NUMBER OF SEQUENCES: 79
15 CORRESPONDENCE ADDRESS:
16 ADDRESSEE: Arnold, White & Durkee
17 STREET: P. O. Box 4433
18 CITY: Houston
19 STATE: Texas
20 COUNTRY: USA
21 ZIP: 77210
22 COMPUTER READABLE FORM:
23 MEDIUM TYPE: Floppy disk
24 COMPUTER: IBM PC compatible
25 OPERATING SYSTEM: PC-DOS/MS-DOS
26 SOFTWARE: PatentIn Release #1.0, Version #1.30
27 CURRENT APPLICATION DATA:
28 APPLICATION NUMBER: US/08/784,582
29 FILING DATE: Concurrently Herewith
30 CLASSIFICATION: 435
31 PRIOR APPLICATION DATA:
32 APPLICATION NUMBER: US 60/028,427
33 FILING DATE: 15-OCT-1996
34 APPLICATION DATA:
35 APPLICATION NUMBER: US 08/589,028
36 FILING DATE: 19-JAN-1996
37 ATTORNEY/AGENT INFORMATION:
38 NAME: Highlander, Steven L.
39 REGISTRATION NUMBER: 37,642
40 REFERENCE/DOCKET NUMBER: UTSD:514
41 TELECOMMUNICATION INFORMATION:
42 TELEPHONE: 512/418-3000
43 TELEFAX: 512/474-7577
44 INFORMATION FOR SEQ ID NO: 2:
45 SEQUENCE CHARACTERISTICS:
46 LENGTH: 110 amino acids
47 TYPE: amino acid
48 STRANDEDNESS:

TOPOLOGY: linear
US-08-784-582-2

Query Match 100.0%; Score 463; DB 3; Length 110;
Best Local Similarity 100.0%; Pred. No. 2,1e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLGGSHLVEALYVCGERGFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
DB 25 FVNHLGGSHLVEALYVCGERGFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84

QY 61 SLQKRGIVEQCCSTCSLYQLNVCN 86
DB 85 SLQKRGIVEQCCSTCSLYQLNVCN 110

RESULT 11
US-08-785-271-2
; Sequence 2, Application US/08785271
; Patent No. 6194176
; GENERAL INFORMATION:
; APPLICANT: Newgard, Christopher B.
; APPLICANT: Halban, Philippe A.
; APPLICANT: No. 6194176minetox, Karl D.
; APPLICANT: Clark, Samuel A.
; APPLICANT: Thigpen, Aimee E.
; APPLICANT: Cuevas, Christian
; APPLICANT: Kruse, Fred
; TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PEPTIDES FROM
; TITLE OF INVENTION: SECRETORY CELL LINES
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Barker
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/785,271
; FILING DATE: Contemporarily Retrieved
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/589,028
; FILING DATE: 19-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven H.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: US/08/513
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 110 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-785-271-2

Query Match 100.0%; Score 463; DB 3; Length 110;
Best Local Similarity 100.0%; Pred. No. 2,1e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLGGSHLVEALYVCGERGFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
DB 25 FVNHLGGSHLVEALYVCGERGFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84

QY 61 SLQKRGIVEQCCSTCSLYQLNVCN 86
DB 85 SLQKRGIVEQCCSTCSLYQLNVCN 110

RESULT 12
US-08-472-701-2
; Sequence 2, Application US/08472701
; Patent No. 6509165
; GENERAL INFORMATION:
; APPLICANT: Griffin, Ann C.
; APPLICANT: Hickey, William F.
; TITLE OF INVENTION: Detection and Treatment Methods for
; TITLE OF INVENTION: Type 1 Diabetes
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,701
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/272,220
; FILING DATE: 08-JULY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Gallo A., Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: DCI-092DV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 110 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-08-472-701-2

Query Match 100.0%; Score 463; DB 4; Length 110;
Best Local Similarity 100.0%; Pred. No. 2,1e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLGGSHLVEALYVCGERGFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
DB 25 FVNHLGGSHLVEALYVCGERGFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84

QY 61 SLQKRGIVEQCCSTCSLYQLNVCN 86
DB 85 SLQKRGIVEQCCSTCSLYQLNVCN 110

RESULT 13
US-09-185-852-2
; Sequence 2, Application US/09185852
; Patent No. 6537806
; GENERAL INFORMATION:
; APPLICANT: Osborne, William R.A.
; APPLICANT: Ramesh, Nagarajan
; TITLE OF INVENTION: Compositions and Methods for Treating Diabetes
; FILE REFERENCE: P-0W 3264
; CURRENT APPLICATION NUMBER: US/09/185,852

QY 61 SLQKRGIVEQCCSTCSLYQLNVCN 86
DB 85 SLQKRGIVEQCCSTCSLYQLNVCN 110

RESULT 12
US-08-472-701-2
; Sequence 2, Application US/08472701
; Patent No. 6509165
; GENERAL INFORMATION:
; APPLICANT: Griffin, Ann C.
; APPLICANT: Hickey, William F.
; TITLE OF INVENTION: Detection and Treatment Methods for
; TITLE OF INVENTION: Type 1 Diabetes
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,701
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/272,220
; FILING DATE: 08-JULY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Gallo A., Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: DCI-092DV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 110 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-08-472-701-2

Query Match 100.0%; Score 463; DB 4; Length 110;
Best Local Similarity 100.0%; Pred. No. 2,1e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLGGSHLVEALYVCGERGFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
DB 25 FVNHLGGSHLVEALYVCGERGFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84

QY 61 SLQKRGIVEQCCSTCSLYQLNVCN 86
DB 85 SLQKRGIVEQCCSTCSLYQLNVCN 110

RESULT 13
US-09-185-852-2
; Sequence 2, Application US/09185852
; Patent No. 6537806
; GENERAL INFORMATION:
; APPLICANT: Osborne, William R.A.
; APPLICANT: Ramesh, Nagarajan
; TITLE OF INVENTION: Compositions and Methods for Treating Diabetes
; FILE REFERENCE: P-0W 3264
; CURRENT APPLICATION NUMBER: US/09/185,852

```

? CURRENT FILING DATE: 1998-11-04
? EARLIER APPLICATION NUMBER: 60/087,660
? EARLIER FILING DATE: 1998-06-02
? NUMBER OF SEQ ID NOS: 11
? SOFTWARE: Patent In Ver. 2.0
? SEQ ID NO 2
? LENGTH: 110
? TYPE: PRI
? ORGANISM: Homo sapiens
US-09-185-852-2

Query Match 100.0%; Score 463; DB 4; Length 110;
Best Local Similarity 100.0%; Pred. No. 2,2e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNOHLCGSHLVALYVCGGFFYTPKTRREAEADLVQGVELGGPGAGSLQPLALEG 86
DB 25 FVNOHLCGSHLVALYVCGGFFYTPKTRREAEADLVQGVELGGPGAGSLQPLALEG 84

QY 61 SLOKRGIVFCCTSGSLYOLENYCN 86
DB 61 SLOKRGIVFCCTSGSLYOLENYCN 86

QY 25 FVNOHLCGSHLVALYVCGGFFYTPKTRREAEADLVQGVELGGPGAGSLQPLALEG 84
DB 25 FVNOHLCGSHLVALYVCGGFFYTPKTRREAEADLVQGVELGGPGAGSLQPLALEG 84

QY 61 SLOKRGIVFCCTSGSLYOLENYCN 86
DB 61 SLOKRGIVFCCTSGSLYOLENYCN 86

QY 85 SLOKRGIVFCCTSGSLYOLENYCN 110
DB 85 SLOKRGIVFCCTSGSLYOLENYCN 110

RESULT 15
US-09-280-030-63
? Sequence 63, Application: US/99280030A
? Patent No. 6506595
? GENERAL INFORMATION:
? APPLICANT: Sato, Seiji
? APPLICANT: Higashikuni, Naohiko
? APPLICANT: Kudo, Toshiyuki
? APPLICANT: Kondo, Masaki
? TITLE OF INVENTION: DNAs ENCODING NEW FUSION PROTEINS AND PROCESSES FOR
? TITLE OF INVENTION: PREPARING USEFUL POLYPEPTIDES THROUGH EXPRESSION OF THE
? FILE REFERENCE: 382,1026
? CURRENT APPLICATION NUMBER: US/99280,030A
? EARLIER FILING DATE: 1999-03-26
? EARLIER FILING DATE: 1999-03-31
? NUMBER OF SEQ ID NOS: 66
? SOFTWARE: Patent In Ver. 2.0
? SEQ ID NO 63
? LENGTH: 117
? TYPE: PRI
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Description of Artificial Sequence: Designated is
? OTHER INFORMATION: an amino acid sequence of
? OTHER INFORMATION: MW2sp-MWpmp10-Met-Proinsulin
US-09-280-030-63

Query Match 100.0%; Score 463; DB 4; Length 117;
Best Local Similarity 100.0%; Pred. No. 2,2e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNOHLCGSHLVALYVCGGFFYTPKTRREAEADLVQGVELGGPGAGSLQPLALEG 60
DB 32 FVNOHLCGSHLVALYVCGGFFYTPKTRREAEADLVQGVELGGPGAGSLQPLALEG 91

QY 61 SLOKRGIVFCCTSGSLYOLENYCN 86
DB 92 SLOKRGIVFCCTSGSLYOLENYCN 117

Search completed: September 15, 2003, 12:05:32
Job time : 20.7276 secs

? CURRENT FILING DATE: 1998-11-04
? EARLIER APPLICATION NUMBER: 60/087,660
? EARLIER FILING DATE: 1998-06-02
? NUMBER OF SEQ ID NOS: 11
? SOFTWARE: Patent In Ver. 2.0
? SEQ ID NO 2
? LENGTH: 110
? TYPE: PRI
? ORGANISM: Homo sapiens
US-09-185-852-2

Query Match 100.0%; Score 463; DB 4; Length 110;
Best Local Similarity 100.0%; Pred. No. 2,2e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNOHLCGSHLVALYVCGGFFYTPKTRREAEADLVQGVELGGPGAGSLQPLALEG 86
DB 25 FVNOHLCGSHLVALYVCGGFFYTPKTRREAEADLVQGVELGGPGAGSLQPLALEG 84

QY 61 SLOKRGIVFCCTSGSLYOLENYCN 86
DB 61 SLOKRGIVFCCTSGSLYOLENYCN 86

QY 85 SLOKRGIVFCCTSGSLYOLENYCN 110
DB 85 SLOKRGIVFCCTSGSLYOLENYCN 110

RESULT 14
PCT-US95-08596-2
? Sequence 2, Application: PCT/US9508596
? GENERAL INFORMATION:
? APPLICANT:
? TITLE OF INVENTION: Proinsulin Peptide Compounds for Diabetes
? TITLE OF INVENTION: and Treating Type 2 Diabetes
? NUMBER OF SEQUENCES: 23
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: LAHIVE & COCKFIELD
? STREET: 50 State Street, Suite 510
? CITY: Boston
? STATE: Massachusetts
? COUNTRY: USA
? ZIP: 02109-1875
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: ASCII Text
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: PCT/US95/08596
? FILING DATE:
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/272,220
? FILING DATE: 08-JULY-1993
? CLASSIFICATION:
? ATTORNEY/AGENT INFORMATION:
? NAME: DeCortio, Giulio A., Jr.
? REGISTRATION NUMBER: 31,103
? REFERENCE/DOCKET NUMBER: PCT-952PC
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (617)227-7400
? TELEFAX: (617)227-5942
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 110 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
PCT-US95-08596-2

Query Match 100.0%; Score 463; DB 5; Length 110;
Best Local Similarity 100.0%; Pred. No. 2,1e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNOHLCGSHLVALYVCGGFFYTPKTRREAEADLVQGVELGGPGAGSLQPLALEG 86
DB 1 FVNOHLCGSHLVALYVCGGFFYTPKTRREAEADLVQGVELGGPGAGSLQPLALEG 86
```

GenCore version 5.1.1.6
Copyright (c) 1995 - 2003 Computer One

OM protein - protein search, using sw model

Run on: September 15, 2003, 12:03:35 : Search time 36.0645 seconds
(without alignments)
347,945 Million cell updates/sec

Title: US-09-423-100-4

Perfect score: 463

Sequence: 1 FVNHLCGSHLVEALYLVCGERGFFYPKTRRREADLQVGVGGPGAGSLQPLALEG 60

Scoring table: BLASTSUM62

Gapop 10.0 , Gapext 0.5

Searched: 541936 seqs, 145912426 residues

Total number of hits satisfying chosen parameters: 541936

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

```
1: /cgn2_6/ptcdat4/1/pubaa/US07_PUBCOMB.pept.*
2: /cgn2_6/ptcdat4/1/pubaa/PCT_NEW_PUB.pept.*
3: /cgn2_6/ptcdat4/1/pubaa/US06_NEW_PUB.pept.*
4: /cgn2_6/ptcdat4/1/pubaa/US06_PUBCOMB.pept.*
5: /cgn2_6/ptcdat4/1/pubaa/US07_NEW_PUB.pept.*
6: /cgn2_6/ptcdat4/1/pubaa/PCTUS_PUBCOMB.pept.*
7: /cgn2_6/ptcdat4/1/pubaa/US09_NEW_PUB.pept.*
8: /cgn2_6/ptcdat4/1/pubaa/US09_PUBCOMB.pept.*
9: /cgn2_6/ptcdat4/1/pubaa/US09A_PUBCOMB.pept.*
10: /cgn2_6/ptcdat4/1/pubaa/US09A_PUBCOMB.pept.*
11: /cgn2_6/ptcdat4/1/pubaa/US09C_PUBCOMB.pept.*
12: /cgn2_6/ptcdat4/1/pubaa/US09C_PUBCOMB.pept.*
13: /cgn2_6/ptcdat4/1/pubaa/US10A_PUBCOMB.pept.*
14: /cgn2_6/ptcdat4/1/pubaa/US10A_PUBCOMB.pept.*
15: /cgn2_6/ptcdat4/1/pubaa/US10C_PUBCOMB.pept.*
16: /cgn2_6/ptcdat4/1/pubaa/US10C_PUBCOMB.pept.*
17: /cgn2_6/ptcdat4/1/pubaa/US06_NEW_PUB.pept.*
18: /cgn2_6/ptcdat4/1/pubaa/US06_PUBCOMB.pept.*
```

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	463	100.0	86	10	US-09-878-380-1
2	463	100.0	86	11	US-09-858-935B-4
3	463	100.0	86	14	US-10-028-410-2
4	463	100.0	86	14	US-10-054-873-4
5	463	100.0	96	10	US-09-947-563-4
6	463	100.0	110	9	US-09-205-658-125
7	463	100.0	110	9	US-09-815-229-3
8	463	100.0	110	10	US-09-804-409A-9
9	463	100.0	110	12	US-09-969-748C-6
10	463	100.0	110	15	US-10-038-686-1
11	463	100.0	110	15	US-10-328-813-2
12	463	100.0	117	9	US-09-280-030-63
13	463	100.0	130	9	US-09-280-030-62
14	457	98.7	96	10	US-09-947-563-5
15	306	66.1	166	9	US-09-925-257-805

15	300	64.8	56	9	US-09-815-229-10	Sequence 10, Appl
16	285	61.6	54	9	US-09-815-229-13	Sequence 13, Appl
17	267	57.7	52	14	US-10-054-873-5	Sequence 5, Appl
18	267	57.7	107	14	US-10-054-873-6	Sequence 6, Appl
19	267	57.7	150	14	US-10-054-873-7	Sequence 7, Appl
20	261.5	56.5	51	11	US-09-858-935B-5	Sequence 5, Appl
21	261.5	56.5	51	14	US-10-028-410-3	Sequence 3, Appl
22	259.5	55.8	124	10	US-09-804-711-12	Sequence 12, Appl
23	258.5	55.8	138	10	US-09-861-687-19	Sequence 19, Appl
24	256.5	55.3	50	14	US-10-066-009A-3	Sequence 3, Appl
25	250.5	54.1	124	9	US-09-736-611-12	Sequence 12, Appl
26	250.5	54.1	125	9	US-09-736-611-10	Sequence 10, Appl
27	250.5	54.1	147	9	US-09-736-611-8	Sequence 8, Appl
28	249.5	53.9	124	9	US-09-740-359-12	Sequence 12, Appl
29	249.5	53.9	124	10	US-09-894-711-12	Sequence 12, Appl
30	249.5	53.9	125	9	US-09-740-359-10	Sequence 10, Appl
31	249.5	53.9	125	10	US-09-894-711-10	Sequence 10, Appl
32	249.5	53.9	147	9	US-09-740-359-7	Sequence 7, Appl
33	249.5	53.9	147	9	US-09-740-359-5	Sequence 5, Appl
34	248	53.6	144	9	US-09-736-611-6	Sequence 6, Appl
35	248	53.6	144	9	US-09-740-359-5	Sequence 5, Appl
36	248	53.6	146	10	US-09-894-711-5	Sequence 5, Appl
37	227.5	49.1	144	10	US-09-894-711-7	Sequence 7, Appl
38	209	45.1	46	9	US-09-205-658-132	Sequence 132, App
39	197	42.5	46	9	US-09-205-658-133	Sequence 133, App
40	184.5	39.8	50	9	US-09-205-658-115	Sequence 115, App
41	180	38.9	46	9	US-09-205-658-135	Sequence 135, App
42	178	38.4	32	9	US-09-815-229-14	Sequence 14, Appl
43	178	38.4	32	10	US-09-947-563-6	Sequence 6, Appl
44	175	37.8	46	9	US-09-205-658-134	Sequence 134, App
45	174	37.6	35	9	US-09-815-229-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-09-878-380-1
Sequence 1, Application US/09/878.380
Patent No. US20020160435A1
GENERAL INFORMATION:
APPLICANT: Fujirebio Inc.
APPLICANT: KITAJIMA, Sachiko
APPLICANT: KURANO, Yoshihiro
APPLICANT: NAKATSUBO, Kaoru
APPLICANT: NISHIZONO, Isao
TITLE OF INVENTION: Immunoassay For Measuring Human C-Peptide and Kit Therefor
FILE REFERENCE: C760-0291P
CURRENT APPLICATION NUMBER: US/09/878.380
PRIOR FILING DATE: 2001-06-12
PRIOR APPLICATION NUMBER: JP 2000-174591
NUMBER OF SEQ IDS NOS: 2
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 1
LENGTH: 86
TYPE: PPT
ORGANISM: Homo sapiens
US-09-878-380-1

Query Match 100.0%; Score 463; DB ID: Length 86.
Best local Similarity 100.0%; Pred. No. 9.7e-45;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q1 1 FVNHLCGSHLVEALYLVCGERGFFYPKTRRREADLQVGVGGPGAGSLQPLALEG 60

D1 1 FVNHLCGSHLVEALYLVCGERGFFYPKTRRREADLQVGVGGPGAGSLQPLALEG 60

Q7 61 SLQKRGIVECCCTSCISLYOLENYCN 86

D7 61 SLQKRGIVECCCTSCISLYOLENYCN 86

RESULT 2

```

US-09-858-935B-4
: Sequence 4, Application US/09H58015R
: Publication No. US2903069177A1
: GENERAL INFORMATION:
: APPLICANT: Dubaqui, Yves
: APPLICANT: Lofman, Henry B.
: TITLE OF INVENTION: METHOD FOR TREATING CARTILAGE DISORDERS
: FILE REFERENCE: P1794R1
: CURRENT APPLICATION NUMBER: US/09/858,935B
: PRIOR FILING DATE: 2002-07-02
: PRIOR APPLICATION NUMBER: US 60/246,985
: PRIOR FILING DATE: 2000-11-15
: PRIOR APPLICATION NUMBER: US 60/204,496
: PRIOR FILING DATE: 2000-05-11
: NUMBER OF SEQ ID NOS: 153
: SEQ ID NO 4
: LENGTH: 86
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-858-935B-4

Query Match      100.0%   Score 463; DB 11; Length 86;
Best Local Similarity 100.0%; Pred. No. 9.7e-45;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQHLCGSHLVEALYLCGERGFFYTKTRREAEDLVQGVLEGGGAGSLQPLALEG 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 FVNQHLCGSHLVEALYLCGERGFFYTKTRREAEDLVQGVLEGGGAGSLQPLALEG 60

QY 61 SLOKRGIVEQCCTSCSLYQLENYCN 86
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 SLOKRGIVEQCCTSCSLYQLENYCN 86

RESULT 3
US-10-028-410-2
: Sequence 2, Application US/10028410
: Publication No. US20020160355A1
: GENERAL INFORMATION:
: APPLICANT: Dubaqui, Yves
: APPLICANT: Lofman, Henry
: TITLE OF INVENTION: PROTEIN VARIANTS
: FILE REFERENCE: P1712R1-1
: CURRENT APPLICATION NUMBER: US/10/028,410
: PRIOR FILING DATE: 2001-12-19
: PRIOR APPLICATION NUMBER: US/09/477,924
: PRIOR FILING DATE: 2000-01-05
: NUMBER OF SEQ ID NOS: 6
: SEQ ID NO 2
: LENGTH: 86
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-028-410-2

Query Match      100.0%   Score 463; DB 14; Length 86;
Best Local Similarity 100.0%; Pred. No. 9.7e-45;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQHLCGSHLVEALYLCGERGFFYTKTRREAEDLVQGVLEGGGAGSLQPLALEG 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 FVNQHLCGSHLVEALYLCGERGFFYTKTRREAEDLVQGVLEGGGAGSLQPLALEG 60

QY 61 SLOKRGIVEQCCTSCSLYQLENYCN 86
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 SLOKRGIVEQCCTSCSLYQLENYCN 86

RESULT 4
US-10-054-873-4
: Sequence 4, Application US/10054873
: Publication No. US20020164712A1
: GENERAL INFORMATION:
: APPLICANT: Dubaqui, Yves
: APPLICANT: Lofman, Henry B.
: TITLE OF INVENTION: METHOD FOR TREATING CARTILAGE DISORDERS
: FILE REFERENCE: P1794R1
: CURRENT APPLICATION NUMBER: US/10/054,873
: PRIOR FILING DATE: 2002-07-02
: PRIOR APPLICATION NUMBER: US 60/246,985
: PRIOR FILING DATE: 2000-11-15
: PRIOR APPLICATION NUMBER: US 60/204,496
: PRIOR FILING DATE: 2000-05-11
: NUMBER OF SEQ ID NOS: 153
: SEQ ID NO 4
: LENGTH: 86
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-054-873-4

Query Match      100.0%   Score 463; DB 14; Length 86;
Best Local Similarity 100.0%; Pred. No. 9.7e-45;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQHLCGSHLVEALYLCGERGFFYTKTRREAEDLVQGVLEGGGAGSLQPLALEG 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 FVNQHLCGSHLVEALYLCGERGFFYTKTRREAEDLVQGVLEGGGAGSLQPLALEG 60

QY 61 SLOKRGIVEQCCTSCSLYQLENYCN 86
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 SLOKRGIVEQCCTSCSLYQLENYCN 86

RESULT 5
US-09-947-563-4
: Sequence 4, Application US/09947563
: Patent No. US20020156234A1
: GENERAL INFORMATION:
: APPLICANT: Rubroder, Franz-Josef
: APPLICANT: Keller, Reinhold
: TITLE OF INVENTION: Improved process for obtaining
: NUMBER OF SEQUENCES: 7 insulin precursors having correctly bonded cystine bri
: CORRESPONDENCE ADDRESS:
: ADDRESS: Finnegan, Henderson, Farrabow, Garrett &
: Duncer
: STREET: 1300 I Street, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20005-3315
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible

```



```
Do 85 SLOKRGIVECCCTSCISLYOLENYCN 110
RESULT 9
US-09-969-748C-6
: Sequence 6, Application US/0969748C
: Publication No. US20030161629A1
: GENERAL INFORMATION:
: APPLICANT: ARZEXE PHARMACEUTICALS, INC.
: APPLICANT: HOUSTON, Lou, L.
: APPLICANT: SHERIDAN, Philip, J.
: APPLICANT: HAWLEY, Stephen
: APPLICANT: GAYN, Jacqueline, M.
: APPLICANT: CHAPIN, Steven
: APPLICANT: BASU, Amaresh
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF SKELETONALLY ACTIVE
: FILE REFERENCE: US/09/969,748C
: CURRENT APPLICATION NUMBER: US/09/969,748C
: PRIOR FILING DATE: 2002-12-10
: PRIOR FILING DATE: 2001-02-06
: PRIOR APPLICATION NUMBER: US 60/267,901
: PRIOR FILING DATE: 2003-11-24
: PRIOR APPLICATION NUMBER: US 60/248,819
: PRIOR FILING DATE: 2003-11-11
: PRIOR APPLICATION NUMBER: US 60/248,478
: PRIOR FILING DATE: 2003-11-11
: PRIOR APPLICATION NUMBER: US 60/247,929
: PRIOR FILING DATE: 2003-10-07
: NUMBER OF SEQ ID NOS: 115
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 6
: LENGTH: 110
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-969-748C-6

Query Match 100.0%; Score 463; DB 12; Length 110;
Best Local Similarity 100.0%; Pred. No. 1.3e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 FVNCHLGGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
Db 25 FVNCHLGGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84

Cy 61 SLOKRGIVECCCTSCISLYOLENYCN 86
Db 85 SLOKRGIVECCCTSCISLYOLENYCN 110

RESULT 10
US-10-038-686-1
: Sequence 1, Application US/10038686
: Publication No. US20030045467A1
: GENERAL INFORMATION:
: APPLICANT: Otsuka, Tibeter
: TITLE OF INVENTION: ANTIOXIDANT VACCINE
: FILE REFERENCE: 10276-067001
: CURRENT APPLICATION NUMBER: US/10/038,686
: CURRENT FILING DATE: 2002-05-31
: PRIOR APPLICATION NUMBER: 60/269,062
: PRIOR FILING DATE: 2001-01-05
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 110
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-038-686-1

Query Match 100.0%; Score 463; DB 15; Length 110;
Best Local Similarity 100.0%; Pred. No. 1.3e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 FVNCHLGGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
Db 25 FVNCHLGGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84

Cy 61 SLOKRGIVECCCTSCISLYOLENYCN 86
Db 85 SLOKRGIVECCCTSCISLYOLENYCN 110

RESULT 11
US-10-328-813-2
: Sequence 2, Application US/10328813
: Publication No. US20030113305A1
: GENERAL INFORMATION:
: APPLICANT: Osborne, William R.A.
: APPLICANT: Ramesh, Nagaratnam
: TITLE OF INVENTION: Compositions and Methods for Treating Diabetes
: FILE REFERENCE: P-CW 3264
: CURRENT APPLICATION NUMBER: US/10/328,813
: CURRENT FILING DATE: 2002-12-23
: PRIOR APPLICATION NUMBER: US/09/185,852
: PRIOR FILING DATE: 1998-11-04
: PRIOR APPLICATION NUMBER: 60/087,660
: PRIOR FILING DATE: 1998-06-02
: NUMBER OF SEQ ID NOS: 11
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 2
: LENGTH: 110
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-328-813-2

Query Match 100.0%; Score 463; DB 15; Length 110;
Best Local Similarity 100.0%; Pred. No. 1.3e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 FVNCHLGGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
Db 25 FVNCHLGGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84

Cy 61 SLOKRGIVECCCTSCISLYOLENYCN 86
Db 85 SLOKRGIVECCCTSCISLYOLENYCN 110

RESULT 12
US-09-280-030-63
: Sequence 63, Application US/09280030A
: Patent No. US2001002155A1
: GENERAL INFORMATION:
: APPLICANT: Sato, Seiji
: APPLICANT: Higashikuni, Naohiko
: APPLICANT: Kudo, Toshiyuki
: APPLICANT: Kondo, Masaaki
: TITLE OF INVENTION: DNAS ENCODING NEW FUSION PROTEINS AND PROCESSES FOR
: TITLE OF INVENTION: PREPARING USEFUL POLYPEPTIDES THROUGH EXPRESSION OF THE
: FILE REFERENCE: DNAS
: FILE REFERENCE: 382.1026
: CURRENT APPLICATION NUMBER: US/09/280,030A
: CURRENT FILING DATE: 1999-03-26
: EARLIER APPLICATION NUMBER: JP10-87339/1998
: EARLIER FILING DATE: 1998-03-31
: NUMBER OF SEQ ID NOS: 66
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 63
: LENGTH: 117
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: Designated is
: OTHER INFORMATION: an amino acid sequence of
: OTHER INFORMATION: MAPsp-Mwmp10-Met-Proinsulin
US-09-280-030-63
```

```

Query Match      100.0% Score 463; DB 9; Length 170;
Best Local Similarity 100.0%; Pred. No. 1.4e-44;
Matches 86; Conservative 0; Mismatches 0; Gaps 0;

QY 1 FVNHLCGSHVEALYVCGEGFFYTPKTRREADLVGVGVGGAGSLQPLALEG 60
DB 32 FVNHLCGSHVEALYVCGEGFFYTPKTRREADLVGVGVGGAGSLQPLALEG 61

QY 61 SLOKRGIVEQCCTSCSLYQLENYC 86
DB 92 SLOKRGIVEQCCTSCSLYQLENYC 117

RESULT 13
US-09-280-030-62
: Sequence 62, Application US/09280030A
: Patent No. US20030021519A1
: GENERAL INFORMATION:
: APPLICANT: Satoh, Seiji
: APPLICANT: Higashikuni, Naohiko
: APPLICANT: Kudo, Toshiyuki
: APPLICANT: Kondo, Masaaki
: TITLE OF INVENTION: DNAs ENCODING NEW FUSION PROTEINS AND PROCESSES FOR
: TITLE OF INVENTION: PREPARING USEFUL POLYPEPTIDES THROUGH EXPRESSION OF THE
: FILE REFERENCE: 382.1026
: CURRENT APPLICATION NUMBER: US/99/285,030A
: EARLIER FILING DATE: 1999-03-26
: EARLIER APPLICATION NUMBER: JP10-87439/-99#
: NUMBER OF SEQ ID NOS: 66
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 62
: LENGTH: 130
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: Postulated is
: OTHER INFORMATION: an amino acid sequence of
: OTHER INFORMATION: MWSP:WNPp10-(His)6-Linker-Met-Proinsulin
US-09-280-030-62

Query Match      100.0% Score 463; DB 9; Length 170;
Best Local Similarity 100.0%; Pred. No. 1.4e-44;
Matches 86; Conservative 0; Mismatches 0; Gaps 0;

QY 1 FVNHLCGSHVEALYVCGEGFFYTPKTRREADLVGVGVGGAGSLQPLALEG 60
DB 45 FVNHLCGSHVEALYVCGEGFFYTPKTRREADLVGVGVGGAGSLQPLALEG 164

QY 61 SLOKRGIVEQCCTSCSLYQLENYC 86
DB 105 SLOKRGIVEQCCTSCSLYQLENYC 130

RESULT 14
US-09-947-563-5
: Sequence 5, Application US/99947563
: Patent No. US20030150234A1
: GENERAL INFORMATION:
: APPLICANT: Rubtsov, Franz-Josef
: APPLICANT: Keller, Reinhold
: TITLE OF INVENTION: Improved process for obtaining
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
: Berner
: STREET: 1300 I Street, N.W.
: City: Washington
: STATE: D.C.
: COUNTRY: USA

```

```

ZIP: 20005-3325
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/947,563
FILING DATE: 07-Sep-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/134,836
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Leslie McDonnell
REGISTRATION NUMBER: 34,972
REFERENCE/DOCKET NUMBER: 02481.1600-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 95 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: ORIGINAL SOURCE:
: ORGANISM: Escherichia coli
: FEATURE:
: NAME/KEY: Protein
: LOCATION: 1..96
: SEQUENCE DESCRIPTION: SEQ ID NO: 5:
DS-09-947-563-5

Query Match      98.7% Score 457; DB 10; Length 96;
Best Local Similarity 100.0%; Pred. No. 5.2e-44;
Matches 85; Conservative 0; Mismatches 0; Gaps 0;

QY 1 FVNHLCGSHVEALYVCGEGFFYTPKTRREADLVGVGVGGAGSLQPLALEG 60
DB 11 FVNHLCGSHVEALYVCGEGFFYTPKTRREADLVGVGVGGAGSLQPLALEG 70

QY 61 SLOKRGIVEQCCTSCSLYQLENYC 85
DB 71 SLOKRGIVEQCCTSCSLYQLENYC 95

RESULT 15
US-09-925-237-805
: Sequence 805, Application US/99925237
: Patent No. US20020081659A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
: FILE REFERENCE: PA105
: CURRENT APPLICATION NUMBER: US/03/925,297
: PRIOR FILING DATE: 2001-08-13
: PRIOR APPLICATION NUMBER: PCI/US00/05489
: PRIOR FILING DATE: 2000-03-08
: PRIOR APPLICATION NUMBER: 60/124,270
: PRIOR FILING DATE: 1999-03-12
: NUMBER OF SEQ ID NOS: 928
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 805
: LENGTH: 166
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (84)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE

```


bioinformatics version 5.1.6
Copyright (c) 1993 - 2003. Copyright Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2003, 11:56:45 / Search time 14.1792 seconds
without alignment
531.264 Million cell updates/sec

Title: US-09-423-100-4
Perfect score: 453
Sequence: 1 FVNHLCGSHLVEALYVWG.....IVFQCCSTSCSYLENVON 36

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 96158682 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR_761*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	453	100.0	110	1	IPRH	insulin precursor
2	463	100.0	110	2	A42179	insulin precursor
3	456	98.5	110	2	B42179	insulin precursor
4	456	98.5	110	2	J06178	insulin precursor
5	424	91.6	110	1	INSH	insulin precursor
6	417	90.1	110	1	IPRG	insulin precursor
7	394	85.1	86	1	IPRG	insulin precursor
8	394	85.1	110	1	IPRG	insulin precursor
9	394	85.1	110	1	INS2	insulin precursor
10	392	84.7	108	2	A39843	insulin precursor
11	392	84.7	110	2	I48156	insulin precursor
12	385	83.2	110	1	IPRT1	insulin precursor
13	383	82.7	84	1	IPRG	insulin precursor
14	366.5	79.2	105	1	IPRG	insulin precursor
15	366	79.0	108	1	INS1	insulin precursor
16	334.5	72.2	108	2	S09274	insulin precursor
17	320.5	69.2	77	1	INSH	insulin precursor
18	314	67.8	110	1	IPRG	insulin precursor
19	277.5	59.9	109	1	IPRG	insulin precursor
20	276.5	59.7	103	2	I51221	insulin precursor
21	265.5	57.3	106	1	IPX12	insulin precursor
22	265.5	57.3	107	1	IPCH	insulin precursor
23	262.5	56.7	106	1	IPX11	insulin precursor
24	256.5	55.4	51	1	IPWH1	insulin - sperm whale
25	256.5	55.4	51	1	INMHF	insulin - finback
26	256.5	55.4	51	1	INEL	insulin - elephant
27	256.5	55.4	81	1	IPDK	insulin precursor
28	256	55.3	96	2	PC7082	epidermal growth factor
29	254.5	55.0	51	1	INRY	insulin - hamster

RESULT 1

IPRH

insulin precursor [validated] - human

N:Alternate names: preproinsulin

C:Species: Homo sapiens (man)

C:Date: 23-Oct-1981 #sequence revision 23-Oct-1981 #text change 08-Dec-2000

C:Accession: A93222; A94253; A93216; A94251; A93144; A92075; A9186; I58114; A0157

R:Reil, G.I.; Pictet, R.L.; Rutter, W.J.; Cordell, B.; Tischer, E.; Goodman, H.M.

Nature 284, 26-32, 1980

A:Title: Sequence of the human insulin gene.

A:Reference number: A93222; MUID:80120725; PMID:6243748

A:Accession: A93222

A:Molecule type: DNA

A:Residues: 1-110 <BEL>

A:Cross-references: GB:J00265; NID:g186429; PIDN:AAAS9172.1; PID:g386828

R:Ulrich, A.; Dull, T.J.; Gray, A.; Brosius, J.; Sures, I.

Science 205, 612-615, 1980

A:Title: Genetic variation in the human insulin gene.

A:Reference number: A94253; MUID:80236313; PMID:6248962

A:Accession: A94253

A:Molecule type: DNA

A:Residues: 1-110 <BEL>

A:Cross-references: GB:J00265; NID:g186429; PIDN:AAAS9172.1; PID:g386828

R:Ulrich, A.; Dull, T.J.; Gray, A.; Brosius, J.; Sures, I.; Rutter, W.J.

Nature 282, 525-527, 1979

A:Title: Nucleotide sequence of a cDNA clone encoding human preproinsulin.

A:Reference number: A93216; MUID:80054779; PMID:501234

A:Accession: A93216

A:Molecule type: mRNA

A:Residues: 1-110 <BEL>

A:Cross-references: GB:J00265; NID:g186429; PIDN:AAAS9172.1; PID:g386828

R:Sures, I.; Soedel, D.V.; Gray, A.; Ulrich, A.

Science 208, 57-59, 1980

A:Title: Nucleotide sequence of human preproinsulin complementary DNA.

A:Reference number: A94251; MUID:80147417; PMID:6927840

A:Accession: A94251

A:Molecule type: mRNA

A:Residues: 1-110 <BEL>

A:Cross-references: GB:J00265; NID:g186429; PIDN:AAAS9172.1; PID:g386828

R:Nicol, D.S.H.W.; Smith, I.F.

Nature 187, 483-485, 1960

A:Title: Amino-acid sequence of human insulin.

A:Reference number: A93144

A:Accession: A93144

A:Molecule type: protein

R:Over, P.F.; Cho, S.; Peterson, J.D.; Steiner, D.F.

J. Biol. Chem. 246, 1375-1386, 1971

A:Title: Studies on human proinsulin. Isolation and amino acid sequence of the human

A:Reference number: A92075; MUID:71116410; PMID:5101771

A:Accession: A92075

A:Molecule type: protein

A:Residues: 57-87 <OE>
 R:Ko, A.; Smyth, D.G.; Markussen, J.; Sundby, F.
 Eur. J. Biochem. 20: 190-199, 1971
 A:Title: Amino acid sequence of the C-peptide of human proinsulin.
 A:Reference number: A91196; MUI:71257722; PMID:560484
 A:Accession: A91186
 A:Molecule type: protein.
 A:Residues: 57-87 <KOA>
 A:Cross-references: EMBL:X61689; NID:938251; PID:CAA43403.1; PID:938252
 A:Note: sequence extracted from NCBI backbone (NCBIP:95067)
 C:Species: chimpanzee
 A:Introns: 63/2
 C:Superfamily: insulin

Query Match 100.0%; Score 463; DB 2; Length 110;
 Best Local Similarity 100.0%; Pred. No. 2.8e-43;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVGAGSLQPLALEG 60
 DQ 25 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVGAGSLQPLALEG 84
 QY 61 SLOKRGIVEQCCTSIQSLYLENYCN 86
 DB 85 SLOKRGIVEQCCTSIQSLYLENYCN 110

RESULT 3
 A42179
 Insulin precursor - green monkey
 C:Species: Cercopithecus aethiops (green monkey, grivet)
 C:Date: 04-Mar-1993; sequence_revision 18-Nov-1994; text_change 16-Jul-1999
 A:Reference number: A42179; A05232; S16494; S22056
 A:Accession: A42179; A05232; S16494; S22056
 R:Seino, S.; Bell, G.I.; Li, W.H.
 Mol. Biol. Evol. 9: 193-203, 1992
 A:Title: Sequences of primate insulin genes support the hypothesis of a slower rate
 A:Reference number: A42179; MUI:92219553; PMID:1560757
 A:Accession: A42179
 A:Molecule type: DNA
 A:Residues: 1-110 <SEI>
 A:Cross-references: EMBL:X61092; NID:922808; PID:CAA43405.1; PID:922809
 A:Note: sequence extracted from NCBI backbone (NCBIN:95185; NCBIP:95194)
 R:Peterson, J.D.; Nehrlich, S.; Oyer, P.E.; Steiner, D.F.
 J. Biol. Chem. 247: 4866-4871, 1972
 A:Title: Determination of the amino acid sequence of the monkey, sheep, and dog pro
 A:Reference number: A52111; MUI:72258016; PMID:4626369
 A:Accession: A05232
 A:Molecule type: protein
 A:Residues: 57-87 <PEP>
 C:Genetics:
 A:Introns: 63/2
 C:Superfamily: insulin
 C:Keywords: hormone; pancreas
 F:1-24/Domain: signal sequence; #status predicted <SIG>
 F:25-54/Domain: insulin chain B; #status predicted <BC>
 F:25-54,90-110/Product: insulin; #status predicted <MAT>
 F:57-87/Domain: connecting peptide; #status experimental <CPEP>
 F:90-110/Domain: insulin chain A; #status predicted <ACH>
 F:31-96,43-109,95-100/Disulfide bonds: #status predicted

Query Match 98.5%; Score 456; DB 2; Length 110;
 Best Local Similarity 98.8%; Pred. No. 1.6e-42;
 Matches 85; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVGAGSLQPLALEG 60
 DB 25 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVGAGSLQPLALEG 84
 QY 61 SLOKRGIVEQCCTSIQSLYLENYCN 86
 DB 85 SLOKRGIVEQCCTSIQSLYLENYCN 110

RESULT 4
 A42179
 Insulin precursor - chimpanzee
 C:Species: Pan troglodytes (chimpanzee)
 C:Date: 04-Mar-1993; sequence_revision 18-Nov-1994; text_change 16-Jul-1999
 A:Reference number: A42179; S22058
 A:Accession: A42179; S22058
 R:Seino, S.; Bell, G.I.; Li, W.H.

A:Residues: 57-87 <OE>
 R:Ko, A.; Smyth, D.G.; Markussen, J.; Sundby, F.
 Eur. J. Biochem. 20: 190-199, 1971
 A:Title: Amino acid sequence of the C-peptide of human proinsulin.
 A:Reference number: A91196; MUI:71257722; PMID:560484
 A:Accession: A91186
 A:Molecule type: protein.
 A:Residues: 57-87 <KOA>
 A:Cross-references: EMBL:X61689; NID:938251; PID:CAA43403.1; PID:938252
 A:Note: sequence extracted from NCBI backbone (NCBIP:95067)
 C:Species: chimpanzee
 A:Introns: 63/2
 C:Superfamily: insulin

Query Match 100.0%; Score 463; DB 1; Length 110;
 Best Local Similarity 100.0%; Pred. No. 2.8e-43;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVGAGSLQPLALEG 60
 DB 25 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVGAGSLQPLALEG 84
 QY 61 SLOKRGIVEQCCTSIQSLYLENYCN 86
 DB 85 SLOKRGIVEQCCTSIQSLYLENYCN 110

RESULT 2
 A42179
 Insulin precursor - chimpanzee
 C:Species: Pan troglodytes (chimpanzee)
 C:Date: 04-Mar-1993; sequence_revision 18-Nov-1994; text_change 16-Jul-1999
 A:Reference number: A42179; S22058
 A:Accession: A42179; S22058
 R:Seino, S.; Bell, G.I.; Li, W.H.

JQ0178

insulin precursor - crab-eating macaque
 C:Species: Macaca fascicularis (crab-eating macaque)
 C:Date: 07-Sep-1990 #sequence_revision 37-Sep-1990 #text_change 16-Jul-1990
 C:Accession: JQ0178
 R:Wetzel, W.; Gronow, J.; Leineweber, M.; Wentzel, F.; Winkler, F.
 Gene 19, 179-183, 1982
 A:Title: The nucleotide sequence of cDNA coding for proinsulin from the primate Macaca fascicularis
 A:Reference number: JQ0178; MUID:830474; PMID:6184262

A:Accession: JQ0178
 A:Molecule type: mRNA
 A:Residues: 1-110 <SWT>
 A:Cross-references: GB:J00346; NID:9342121; PIDN:AAA664311; PMID:6184262
 C:Superfamily: insulin
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-54/Domain: insulin #status predicted <MAT>
 F:55-89/Domain: insulin chain B #status predicted <SCH>
 F:90-110/Domain: insulin connecting C peptide #status predicted <CPEP>
 F:91-96/Domain: insulin chain A #status predicted <ACH>
 F:91-96,43-109,95-100/Disulfide bonds: #status predicted

Query Match 98.5% Score 456; DB 2; Length 110;

Best Local Similarity 98.8% Pred. No. 1.6e-42; Mismatches 1; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAFLDQVGVGVGGGSGSLQPLALEG 60

DB 25 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAFLDQVGVGVGGGSGSLQPLALEG 84

QY 61 SLQKRGIVEQCCTISCSLYOLENYCN 86

DB 95 SLQKRGIVEQCCTISCSLYOLENYCN 110

RESULT 5

insulin precursor - rabbit
 K:Alternate names: proinsulin
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 24-Apr-1984 #sequence_revision 25-Aug-1997 #text_change 18-Jun-1999
 C:Accession: A53438; A01581
 R:Devaskar, S.U.; Giddings, S.J.; Rajakumar, P.A.; Carnaghi, L.F.; Menon, R.K.; Zahm, E.
 J. Biol. Chem. 269, 8445-8454, 1994
 A:Title: Insulin gene expression and insulin synthesis in mammalian neurons: implications for the pathogenesis of diabetes mellitus
 A:Reference number: A53438; MUID:94170230; PMID:814250
 A:Accession: A53438
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-110 <DEV>
 A:Cross-references: GB:003610; NID:9407970; PIDN:AAA196311; PMID:814250
 R:Smith, L.F.

Query Match 91.6% Score 424; DB 1; Length 110;

Best Local Similarity 90.7% Pred. No. 5e-39;

Matches 78; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAFLDQVGVGVGGGSGSLQPLALEG 60

DB 25 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAFLDQVGVGVGGGSGSLQPLALEG 84

QY 61 SLQKRGIVEQCCTISCSLYOLENYCN 86

DB 95 SLQKRGIVEQCCTISCSLYOLENYCN 110

QY 61 SLQKRGIVEQCCTISCSLYOLENYCN 86

DB 95 SLQKRGIVEQCCTISCSLYOLENYCN 110

QY 61 SLQKRGIVEQCCTISCSLYOLENYCN 86

DB 95 SLQKRGIVEQCCTISCSLYOLENYCN 110

QY 61 SLQKRGIVEQCCTISCSLYOLENYCN 86

DB 95 SLQKRGIVEQCCTISCSLYOLENYCN 110

QY 61 SLQKRGIVEQCCTISCSLYOLENYCN 86

DB 95 SLQKRGIVEQCCTISCSLYOLENYCN 110

QY 61 SLQKRGIVEQCCTISCSLYOLENYCN 86

DB 95 SLQKRGIVEQCCTISCSLYOLENYCN 110

QY 61 SLQKRGIVEQCCTISCSLYOLENYCN 86

DB 95 SLQKRGIVEQCCTISCSLYOLENYCN 110

QY 61 SLQKRGIVEQCCTISCSLYOLENYCN 86

DB 95 SLQKRGIVEQCCTISCSLYOLENYCN 110

QY 61 SLQKRGIVEQCCTISCSLYOLENYCN 86

DB 95 SLQKRGIVEQCCTISCSLYOLENYCN 110

QY 61 SLQKRGIVEQCCTISCSLYOLENYCN 86

DB 95 SLQKRGIVEQCCTISCSLYOLENYCN 110

QY 61 SLQKRGIVEQCCTISCSLYOLENYCN 86

DB 95 SLQKRGIVEQCCTISCSLYOLENYCN 110

RESULT 6

IPH6

insulin precursor - dog

C:Species: Canis lupus familiaris (dog)
 C:Date: 24-Apr-1984 #sequence_revision 15-Nov-1984 #text_change 16-Jul-1999
 C:Accession: A92413; A01587; S16493
 R:Kawox, S.C.M.; Chan, S.J.; Steiner, D.F.
 J. Biol. Chem. 258, 2357-2363, 1983

A:Title: Cloning and nucleotide sequence analysis of the dog insulin gene. Coded at the 5' end of the insulin gene
 A:Reference number: A92413; MUID:83109071; PMID:6296142

A:Accession: A92413

A:Molecule type: DNA

A:Residues: 1-110 <SMI>

A:Cross-references: GB:V00179; GB:J00042; NID:9994; PIDN:CAA23475.1; PID:9995

R:Smith, L.F.

Am. J. Med. 40, 662-666, 1966

A:Title: Species variation in the amino acid sequence of insulin.

A:Reference number: A90029; MUID:66160119; PMID:5949593

A:Accession: A01587

A:Molecule type: protein

A:Residues: 25-54; 90-110 <SMI>

F:Peterson, J.D.; Nehrlich, S.; Over, P.E.; Steiner, D.F.

J. Biol. Chem. 247, 4866-4871, 1972

A:Title: Determination of the amino acid sequence of the monkey, sheep, and dog pr

A:Reference number: A92111; MUID:72258016; PMID:4626369

A:Accession: S16493

A:Molecule type: protein

A:Residues: 65-85; 117-87 <PEP>

C:Superfamily: insulin

C:Keywords: hormone; pancreas

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-54/Domain: insulin chain B #status experimental <BCH>

F:55-89/Domain: insulin chain A #status experimental <MAT>

F:90-110/Domain: connecting peptide #status predicted <CPEP>

F:91-96/Domain: insulin chain A #status experimental <ACH>

F:91-96,43-109,95-100/Disulfide bonds: #status experimental

Query Match 90.1% Score 417; DB 1; Length 110;

Best Local Similarity 89.5% Pred. No. 2.9e-38;

Matches 77; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAFLDQVGVGVGGGSGSLQPLALEG 60

DB 25 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAFLDQVGVGVGGGSGSLQPLALEG 84

QY 61 SLQKRGIVEQCCTISCSLYOLENYCN 86

DB 95 SLQKRGIVEQCCTISCSLYOLENYCN 110

QY 61 SLQKRGIVEQCCTISCSLYOLENYCN 86

DB 95 SLQKRGIVEQCCTISCSLYOLENYCN 110

QY 61 SLQKRGIVEQCCTISCSLYOLENYCN 86

DB 95 SLQKRGIVEQCCTISCSLYOLENYCN 110

QY 61 SLQKRGIVEQCCTISCSLYOLENYCN 86

DB 95 SLQKRGIVEQCCTISCSLYOLENYCN 110

QY 61 SLQKRGIVEQCCTISCSLYOLENYCN 86

DB 95 SLQKRGIVEQCCTISCSLYOLENYCN 110

QY 61 SLQKRGIVEQCCTISCSLYOLENYCN 86

DB 95 SLQKRGIVEQCCTISCSLYOLENYCN 110

QY 61 SLQKRGIVEQCCTISCSLYOLENYCN 86

DB 95 SLQKRGIVEQCCTISCSLYOLENYCN 110

QY 61 SLQKRGIVEQCCTISCSLYOLENYCN 86

DB 95 SLQKRGIVEQCCTISCSLYOLENYCN 110

QY 61 SLQKRGIVEQCCTISCSLYOLENYCN 86

DB 95 SLQKRGIVEQCCTISCSLYOLENYCN 110

QY 61 SLQKRGIVEQCCTISCSLYOLENYCN 86

DB 95 SLQKRGIVEQCCTISCSLYOLENYCN 110

QY 61 SLQKRGIVEQCCTISCSLYOLENYCN 86

DB 95 SLQKRGIVEQCCTISCSLYOLENYCN 110

QY 61 SLQKRGIVEQCCTISCSLYOLENYCN 86

DB 95 SLQKRGIVEQCCTISCSLYOLENYCN 110

A:Accession: A92120
A:Molecule type: protein
A:Residues: 33-63 <TAG>
C:Comment: X's at positions 31-33 and 64-65 represent paired basic residues assumed (by
C:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-30/Domain: insulin chain B #status experimental <BCH>
F:33-63/Domain: connecting peptide #status experimental <CPEP>
F:66-86/Domain: insulin chain A #status experimental <ACH>
F:7-72,79-85,71-75/Disulfide bonds: #status predicted

Query Match 85.1% Score 394; DB 1; Length 86;
Best Local Similarity 84.9% Pred. No. 9.2e-36;
Matches 73; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLEG 60
DB 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLEG 60
QY 61 SLQKRGIVEQCCTSCSLYLENYCN 86
DB 61 FQXXGIVEQCCTSCSLYLENYCN 86

RESULT 8
IPT2
Insulin 2 precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 23-Oct-1981 #sequence_revision 23-Oct-1981 #text_change 21-Sep-1999
C:Accession: B90789; B94231; C92120; I64863; A01570; B42120
R:Lomedico, P.; Rosenthal, A.; Elstratadis, A.; Gilbert, W.; Klotzel, H.; R
Cell 18, 545-558, 1979
A:Title: The structure and evolution of the two nonallelic rat preproinsulin genes.
A:Reference number: A90789; MUID:8004535; PMID:543284
A:Accession: B90789
A:Molecule type: DNA
A:Residues: 1-110 <LOM>
A:Cross-references: G5300748; NID:920458; PIDN:AAA41443.1; PID:920458
R:Steiner, D.F.; Clark, J.L.; Nolan, C.; Rubenstein, A.H.; Matquiesh, E.; Arai, H.; Oy
Recent Prog. Horm. Res. 25, 207-282, 1969
A:Title: Proinsulin and the biosynthesis of insulin.
A:Reference number: A94231; MUID:7096763; PMID:4311936
A:Accession: B94231
A:Molecule type: protein
A:Residues: 25-54; 90-110 <STE>
R:Tager, H.S.; Steiner, D.F.
J. Biol. Chem. 247, 7936-7940, 1972
A:Title: Primary structures of the proinsulin connecting peptides of the rat and horse.
A:Reference number: A92120; MUID:73061498; PMID:4645931
A:Accession: C92120
A:Molecule type: protein
A:Residues: 57-87 <TAG>
R:Lomedico, P.T.; Rosenthal, N.; Klotzel, H.; Elstratadis, A.; Gilbert, W.
Ann. N.Y. Acad. Sci. 343, 425-432, 1980
A:Title: The structure of rat preproinsulin genes.
A:Reference number: I51945; MUID:80240379; PMID:6245167
A:Accession: I64880
A:Status: preliminary; translated from GR/EMBL/DBTJ
A:Molecule type: DNA
A:Residues: 1-110 <RES>
A:Cross-references: GR:M25595; NID:9204550; PIDN:AAA41446.1; PID:9204552
C:Genetics:
A:Gene: INS2
A:Introns: 63/1
C:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-54/Domain: insulin chain B #status experimental <BCH>
F:57-87/Domain: connecting peptide #status experimental <CPEP>
F:90-110/Domain: insulin chain A #status experimental <ACH>
F:31-96,43-109,95-100/Disulfide bonds: #status experimental

Query Match 85.1% Score 394; DB 1; Length 110;
Best Local Similarity 84.9% Pred. No. 9.2e-36;
Matches 73; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLEG 60
DB 25 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLEG 84
QY 61 SLQKRGIVEQCCTSCSLYLENYCN 86
DB 85 AROKRGIVDQCCTSCSLYLENYCN 110

RESULT 9
LNK52
Insulin 2 precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 31-Mar-1992 #sequence_revision 14-Jul-1994 #text_change 18-Jun-1999
C:Accession: A26342; B48172; A61012; B01592
R:Kotwirth, H.M.; Schaefer, I.M.; Villa-Komaroff, L.; Chirgwin, J.M.
J. Mol. Evol. 23, 305-312, 1986
A:Title: Characterization of the two nonallelic genes encoding mouse preproinsulin.
A:Reference number: A92965; MUID:87169768; PMID:3104603
A:Accession: A26342
A:Molecule type: DNA
A:Residues: 1-110 <MEN>
A:Cross-references: GB:X04724; NID:952714; PIDN:CAA28433.1; PID:952715
R:Swia, T.; Ogata, S.; Morioka, H.; Yano, S.
J. Mol. Endocrinol. 5, 61-67, 1990
A:Title: Molecular cloning and DNA sequence analysis of preproinsulin genes in the
A:Reference number: A48172; MUID:90372989; PMID:2377023
A:Accession: B48172
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-110 <SAW>
R:Linde, S.; Nielsen, J.H.; Hansen, B.; Wellinder, B.S.
J. Chromatogr. 462, 243-254, 1989
A:Title: Reversed-phase high-performance liquid chromatographic analyses of insulin
A:Reference number: A61012; MUID:89292078; PMID:2661585
A:Accession: A61012
A:Molecule type: protein
A:Residues: 57-87 <LIN>
R:Buenzli, H.F.; Glatthaar, B.; Kienz, P.; Muelhaupt, E.; Humbel, R.E.
Hoppe-Seyler's Z. Physiol. Chem. 353, 451-458, 1972
A:Title: Amino acid sequence of the two insulins from mouse (Mus musculus).
A:Reference number: A01592; MUID:72189455; PMID:5063718
A:Accession: B01592
A:Molecule type: protein
A:Residues: 25-54; 90-110 <BOR>
C:Genetics:
A:Introns: 63/1
C:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-54/Domain: insulin chain B #status experimental <BCH>
F:57-87/Domain: connecting peptide #status experimental <CPEP>
F:90-110/Domain: insulin chain A #status experimental <ACH>
F:31-96,43-109,95-100/Disulfide bonds: #status predicted

Query Match 85.1% Score 394; DB 1; Length 110;
Best Local Similarity 84.9% Pred. No. 9.2e-36;
Matches 73; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLEG 60
DB 25 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLEG 84
QY 61 SLQKRGIVEQCCTSCSLYLENYCN 86
DB 85 AROKRGIVDQCCTSCSLYLENYCN 110

```

RESULT 10
A39883
insulin precursor - euro-sculi
C:Species: Aotus trivirgatus (dormouse), little monkey, owl monkey
C:Date: 27-Nov-1991; #sequence_revision 27; E.A.:1991; #text_change 16-Jul-1999
C:Accession: A39883
R:Seino, S.; Steiner, D.F.; Bell, G.I.
Proc. Natl. Acad. Sci. U.S.A. 84: 7423-7427, 1987
A:Title: Sequence of a New World primate insulin having low biological potency and immunoreactivity
A:Reference number: A39883; MUID:88041119; PMID:3114367
A:Accession: A39883
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-108 <SEI>
A:Cross-references: GR:J02969; NID:q174555; PIDN:AAA15374.1; PID:q174555
C:Superfamily: insulin

Query Match: 84.7%; Score 392; DB 2; Length 108;
Best Local Similarity 84.9%; Pred. No. 1.5e-35;
Matches 73; Conservative 4; Mismatches 7; Indels 2; Gaps 1;

QY 1 FVNQHLGGSHLVNLYVCCGGRFFYTPKTRREACLOVGVHVGCGGAGSLOPLALEG 60
DB 25 FVNQHLGGSHLVNLYVCCGGRFFYTPKTRREACLOVGVHVGCGGAGSLOPLALEG 60
QY 61 SLQKRGIVDCCCTSCSYQLONYCN 86
DB 83 PMQKRGVDDCCCTSCSYQLONYCN 108

RESULT 11
149166
insulin precursor - golden hamster
C:Species: Mesocricetus auratus (golden hamster)
C:Date: 02-Jun-1996; #sequence_revision 02-Jun-1996; #text_change 16-Jul-1999
C:Accession: 149166
R:Bell, G.I.; Sanchez-Pescador, R.
Diabetes 33: 297-300, 1984
A:Title: Sequence of a cDNA encoding Syrian hamster preproinsulin
A:Reference number: 149166; MUID:84133036; PMID:6365563
A:Accession: 149166
A:Status: preliminary; translated from JB/EMS-1/1842
A:Molecule type: mRNA
A:Residues: 1-110 <RES>
A:Cross-references: GR:M26328; NID:q194240; PIDN:AAA37089.1; PID:q37089
C:Superfamily: insulin

Query Match: 84.7%; Score 372; DB 2; Length 110;
Best Local Similarity 84.9%; Pred. No. 1.5e-35;
Matches 73; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 FVNQHLGGSHLVNLYVCCGGRFFYTPKTRREACLOVGVHVGCGGAGSLOPLALEG 60
DB 25 FVNQHLGGSHLVNLYVCCGGRFFYTPKTRREACLOVGVHVGCGGAGSLOPLALEG 60
QY 61 SLQKRGIVDCCCTSCSYQLONYCN 86
DB 85 AOKKRGIVDCCCTSCSYQLONYCN 110

RESULT 12
149167
insulin 1 precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 23-Oct-1981; #sequence_revision 23-Oct-1981; #text_change 24-Sep-1999
C:Accession: A90788; A90789; A94231; B92120; 151945; A01589
R:Cordell, B.; Bell, G.; Tischer, E.; Donato, F.M.; Ellis, A.; Pirot, P.; Butler, W.
Cell 18: 533-543, 1979
A:Title: Isolation and characterization of a cloned rat insulin gene
A:Reference number: A90788; MUID:80045034; PMID:498283
A:Accession: A90788
A:Molecule type: DNA

```

```

A:Residues: 1-110 <CCR>
A:Cross-references: GR:J00747; NID:q204956; PIDN:AAA1442.1; PID:q204957
R:Comedon, P.; Rosenthal, A.; Efstratiadis, A.; Gilbert, W.; Kolodner, R.; Lizard
Cell 18: 545-558, 1979
A:Title: The structure and evolution of the two nonallelic rat preproinsulin genes
A:Reference number: A90789; MUID:80045035; PMID:498284
A:Accession: A90789
A:Molecule type: DNA
A:Residues: 1-110 <DOM>
A:Cross-references: GR:J00747; NID:q204956; PIDN:AAA1442.1; PID:q204957
R:Steiner, D.F.; Clark, J.L.; Nolan, C.; Rubenstein, A.H.; Margolis, E.; Aten, B
Recent Prog. Horm. Res. 25: 207-282, 1969
A:Title: Proinsulin and the biosynthesis of insulin
A:Reference number: A94231; MUID:70067613; PMID:4311936
A:Accession: A94231
A:Molecule type: protein
A:Residues: 25-54:90-110 <STE>
R:Tager, H.S.; Steiner, D.F.
J. Biol. Chem. 247: 7936-7940, 1972
A:Title: Primary structures of the proinsulin connecting peptides of the rat and h
A:Reference number: A92120; MUID:73061498; PMID:464093;
A:Accession: B92120
A:Molecule type: protein
A:Residues: 57-87 <TAG>
R:Comedon, P.; Rosenthal, A.; Kolodner, R.; Efstratiadis, A.; Gilbert, W.
Ann. N.Y. Acad. Sci. 343: 425-432, 1980
A:Title: The structure of rat preproinsulin genes
A:Reference number: 151945; MUID:80240379; PMID:6249167
A:Accession: 151945
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-110 <RES>
A:Cross-references: GR:M25584; NID:q204947; PIDN:AAA1439.1; PID:q204948
C:Genetics:
A:Gene: INS1
C:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-54/Domain: insulin chain B #status experimental <RCH>
F:25-54:90-110/Product: insulin #status experimental <MAT>
F:57-87/Domain: connecting peptide #status experimental <CPSP>
F:90-110/Domain: insulin chain A #status experimental <ACH>
F:31-95:43-109:95-100/Disulfide bonds: #status experimental

Query Match: 84.2%; Score 385; DB 1; Length 110;
Best Local Similarity 83.7%; Pred. No. 8.8e-35;
Matches 72; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 1 FVNQHLGGSHLVNLYVCCGGRFFYTPKTRREACLOVGVHVGCGGAGSLOPLALEG 60
DB 25 FVNQHLGGSHLVNLYVCCGGRFFYTPKTRREACLOVGVHVGCGGAGSLOPLALEG 60
QY 61 SLQKRGIVDCCCTSCSYQLONYCN 86
DB 85 AOKKRGIVDCCCTSCSYQLONYCN 110

RESULT 13
149166
insulin precursor - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 22-Jun-1981; #sequence_revision 22-Jun-1981; #text_change 16-Jul-1999
C:Accession: A01583; A94572; S16492; A60835; B60835
R:Chance, R.E.; Ellis, R.M.; Bromer, W.W.
Science 161: 165-167, 1968
A:Title: Porcine proinsulin: characterization and amino acid sequence
A:Reference number: A94240; MUID:68286485; PMID:5657063
A:Accession: A01583
A:Molecule type: protein
A:Residues: 1-34; 70-36-84 <CHA>
R:Chance, R.E.
submitted to the Atlas, July 1970
A:Reference number: A94572

```


A:Accession: A94572
A:Molecule type: protein
A:Residues: 1-84 <CH2>
R:Brown, H.; Sanger, F.; Kitai, R.
Biochem. J. 60, 556-565, 1955
A:Title: The structure of pig and sheep insulins.
A:Reference number: A90344
A:Accession: S16492
A:Molecule type: protein
A:Residues: 1-30;31-51 <BRO>
R:Snell, L.; Damgaard, U.
Horm. Metab. Res. 20, 476-480, 1988
A:Title: Proinsulin heterogeneity in pigs.
A:Reference number: A60835; MUID:89032178; PMID:3141865
A:Accession: A60835
A:Molecule type: protein
A:Residues: 33-62 <SN2>
A:Note: the authors report the characterization of a connecting peptide variant lacking
A:Accession: B60835
A:Molecule type: protein
A:Residues: 33-62 <SN2>
R:Blundell, T.L.; Dodson, G.; Hodgkin, D.; Moras, D.
Adv. Protein Chem. 26, 279-402, 1972
A:Title: Insulin: the structure in the crystal and its reflection in chemistry and biology
A:Reference number: A90017
C:Contents: annotation: X-ray crystallography, 1.9 Angstroms
C:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-30/Domain: insulin chain B #status experimental <BCH>
F:1-30,64-84/Product: insulin #status experimental <BCH>
F:33-63/Domain: connecting peptide #status experimental <CPEP>
F:64-84/Domain: insulin chain A #status experimental <ACH>
F:7-70,19-83,69-74/Disulfide bonds: #status experimental

Query Match 82.7% Score 383; DB 1; Length 84;
Best Local Similarity 86.0%; Pred. No. 1,10-14;
Matches 74; Conservative 1; Mismatches 9; Indels 2; Gaps 11

OY : FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQLQSGNGLVGGPGAGSLQPLEAG 86
|||||
DB 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQLQSGNGLVGGPGAGSLQPLEAG 84
|||||

OY 61 SLQKRGIVEQCCTSTCSLYOLENYCN 86
|||||
DB 59 PPQKRGIVEQCCTSTCSLYOLENYCN 84
|||||

RESULT 14
IPBO
Insulin precursor - bovine
C:Species: Bos primigenius taurus (cattle)
C>Date: 24-Apr-1984 #sequence_revision 22-Apr-1995 #text_change 16-Jul-1999
C:Accession: A40909; A92080; A92074; A91185; A90342; A90341; S48185; S46248; A01
R:D'Agostino, J.; Younes, M.A.; White, J.W.; Bosch, P.K.; Field, J.S.; Frazier, M.L.
Mol. Endocrinol. 1, 327-331, 1987
A:Title: Cloning and nucleotide sequence analysis of complementary deoxyribonucleic acid
A:Reference number: A40909; MUID:88288209; PMID:2456452
A:Accession: A40909
A:Molecule type: mRNA
A:Residues: 1-105 <DAA>
A:Cross-references: GB:M54573; MUID:q163573; PDB:AAA30722.1; PDB:q163573
A:Experimental source: fetal pancreas
R:Nolan, C.; Margoliash, E.; Peterson, J.D.; Steiner, D.F.
J. Biol. Chem. 246, 2780-2795, 1971
A:Title: The structure of bovine proinsulin.
A:Reference number: A92080; MUID:71166442; PMID:4928932
A:Accession: A92080
A:Molecule type: protein
A:Residues: 25-105 <NCL>
R:Steiner, D.F.; Cho, S.; Oyer, P.E.; Terris, S.; Peterson, J.D.; Kraustein, A.H.
J. Biol. Chem. 245, 1365-1374, 1971
A:Title: Isolation and characterization of proinsulin C-peptide from bovine pancreas.
A:Reference number: A92074; MUID:71166409; PMID:5545080

A:Accession: A92074
A:Molecule type: protein
A:Residues: 57-82 <SIE>
R:Salokangas, A.; Smyth, D.G.; Markussen, J.; Sundby, F.
Eur. J. Biochem. 20, 183-189, 1971
A:Title: Bovine proinsulin: amino acid sequence of the C-peptide isolated from pancre
A:Reference number: A91185; MUID:71257721; PMID:5105368
A:Accession: A91185
A:Molecule type: protein
A:Residues: 57-82 <SAL>
R:Sanger, F.; Thompson, E.G.P.
Biochem. J. 53, 366-374, 1953
A:Title: The amino-acid sequence in the glycol chain of insulin. 2. The investigati
A:Reference number: A90342
A:Accession: A90342
A:Molecule type: protein
A:Residues: 85-105 <SAN>
R:Sanger, F.; Tuppy, H.
Biochem. J. 49, 481-490, 1951
A:Title: The amino-acid sequence in the phenylalanyl chain of insulin. 2. The inves
A:Reference number: A90341
A:Accession: A90341
A:Molecule type: protein
A:Residues: 25-54 <SA2>
R:Cheng, R.; Kawakishi, S.
Eur. J. Biochem. 223, 759-764, 1994
A:Title: Site-specific oxidation of histidine residues in glycosylated insulin mediated
A:Reference number: S48184; MUID:94333378; PMID:8055951
A:Accession: S48184
A:Molecule type: protein
A:Residues: 85-105 <CHE>
A:Accession: S48195
A>Status: preliminary
A:Molecule type: protein
A:Residues: 25-30; X',32-42; X',44-54 <CH2>
R:Kryte, A.P.; Sanger, F.; Smith, L.F.; Kitai, R.
Biochem. J. 60, 541-556, 1955
A:Title: The disulphide bonds of insulin.
A:Reference number: A90343
C:Contents: annotation: acids; disulfides
R:Kordecki, T.; Eckerskorn, C.; Lottspeich, F.; Baummeister, W.
FEBS Lett. 349, 205-209, 1994
A:Title: Existence of a molecular ruler in proteasomes suggested by analysis of deg
A:Reference number: S46258; MUID:94326921; PMID:8050567
A:Accession: S46258
A>Status: preliminary
A:Molecule type: protein
A:Residues: 25-54 <WEN>
C:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-54/Domain: insulin chain B #status experimental <BCH>
F:25-54,95-105/Product: insulin #status experimental <ACH>
F:57-82/Domain: connecting peptide #status experimental <CPEP>
F:85-105/Domain: insulin chain A #status experimental <ACH>
F:31-91,43-104,90-95/Disulfide bonds: #status experimental

Query Match 79.2% Score 366.5; DB 1; Length 105;
Best Local Similarity 80.2%; Pred. No. 8,7e-33;
Matches 69; Conservative 2; Mismatches 10; Indels 5; Gaps 1;
OY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQLQSGNGLVGGPGAGSLQPLEAG 60
|||||
DB 25 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQLQSGNGLVGGPGAGSLQPLEAG 79
|||||

OY 61 SLQKRGIVEQCCTSTCSLYOLENYCN 86
|||||
DB 80 PPQKRGIVEQCCTSTCSLYOLENYCN 105
|||||

RESULT 15
IPMSI
Insulin 1 precursor - mouse

Tue Sep 16 08:22:52 2003

C:Species: Mus musculus (house mouse)
 C:Date: 24-Apr-1984 #sequence_revision 11-Jul-1994 #text_change 12-Jul-1994
 C:Accession: 926342; A48172; A0_592; B61012
 R:Wentworth, B.M.; Schaefer, S.M.; Villa-Komaroff, L.; Chirgwin, J.M.
 J. Mol. Evol. 23, 305-312, 1986
 A:Title: Characterization of the two nonallelic genes encoding mouse preproinsulin.
 A:Reference number: A92965; MUID:R716975; PMID:3704833
 A:Accession: 926342
 A:Molecule type: DNA
 A:Residues: 1-108 <SEN>
 A:Cross-references: GH:XG4725; NID:952712; PIDN:CAA28434.1; PID:452713
 R:Sawa, T.; Otsuka, S.; Morioka, H.; Yano, S.
 J. Mol. Endocrinol. 5, 61-67, 1990
 A:Title: Molecular cloning and DNA sequence analysis of preproinsulin genes in the M.K.M.
 A:Reference number: A48172; MUID:90372989; PMID:2397023
 A:Accession: A48172
 A>Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-108 <SAW>
 R:Buenzli, H.F.; Glatthard, R.; Kutz, P.; Muelhaupt, E.; Humbel, R.F.;
 Hoppe-Seyler's Z. Physiol. Chem. 353, 451-458, 1972
 A:Title: Amino acid sequence of the two insulins from mouse (Mus musculus).
 A:Reference number: A01592; MUID:72189455; PMID:5063713
 A:Accession: A01592
 A:Molecule type: protein
 A:Residues: 25-54:88-108 <BOE>
 R:Linde, S.; Nielsen, J.H.; Hansen, B.; Welinder, B.S.
 J. Chromatogr. 462, 243-254, 1989.
 A:Title: Reversed-phase high-performance liquid chromatographic analyses of insulin bind
 A:Reference number: A61012; MUID:89292078; PMID:2651585
 A:Accession: B61012
 A:Molecule type: protein
 A:Residues: 57-85 <Lin>
 C:Superfamily: insulin
 C:Keywords: hormone; pancreas
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-54/Domain: insulin chain B #status experimental <BCB>
 F:25-54, 88-108/Product: insulin #status experimental <MAI>
 F:57-85/Domain: connecting pept.de #status experimental <CFEP>
 F:88-108/Domain: insulin chain A #status experimental <ACH>
 F:31-94, 43-107, 93-98/Disulfide bonds: #status predicted

Query Match 79.9% Score 366; DB 1; Length 108;
 Best Local Similarity 81.4% Prod. No. 1212;
 Matches 70; Conservative 4; Mismatches 10; Indels 2; Gaps 1;
 QY 1 FVNCILGSGSLVEALYVPCGRGFFTPKTRPRRRLGQVGVGVGGGPRRRLGSLALR 60
 Db 25 FVKGELGSPHIVKALYLCGRGFFTPKSRREVEVQVELELGGSP--GLQQLALRV 82
 QY 60 SLQKRGIVEGCTSIQS YQDENVCN R6
 Db 83 ARQKRGIVDQCTSIQS YQDENVCN 108

Search completed: September 15, 2003, 12:04:19
 Job time : 14.3792 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Computer Sci.

OM protein - protein search, us sw model

Run on: September 15, 2003, 11:54:00 : Search time 3:01:34 Seconds
(without alignments)
364,433 Million cell updates/sec

Title: US-09-423-100-4

Perfect score: 463

Sequence: 1 FVQVHLCGSHVEALRWG.....IVEGCTTSCSYVQLRYRHN 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47626705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 3

Maximum DB seq length: 260000000

Post-processing: Minimum Match 1%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.1

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %	Score	Match	Length	DB ID	Description
1	463	100.0	110	1	INS_HUMAN	P01309 homo sapien
2	463	100.0	110	1	INS_PANTR	P00110 pan troglod
3	456	98.5	110	1	INS_GORAE	P00478 gorilla
4	456	98.5	110	1	INS_MACFA	P00456 macaca fasc
5	424	91.6	110	1	INS_RABIT	P01311 caprogladus
6	417	90.1	110	1	INS_CANFA	P01321 canis fam
7	413	89.2	110	1	INS_SPERF	P01313 spermophila
8	394	85.1	86	1	INS_BORSE	P01310 bos tauru
9	394	85.1	110	1	INS2_MOUSE	P01326 mus muscul
10	394	85.1	110	1	INS2_RAT	P01335 rattus norv
11	392	84.7	108	1	INS2_COTR	P12604 actus triu
12	392	84.7	110	1	INS_CK110	P01312 cytodidius
13	385	83.2	110	1	INS1_RAT	P01322 rattus norv
14	383	82.7	108	1	INS_PIG	P01315 sus scrofa
15	377	81.4	110	1	INS_PSAAR	P02567 psammomy
16	366.5	79.2	105	1	INS_BOVIN	P01331 bos tauru
17	366	79.0	108	1	INS1_MOUSE	P01325 mus muscul
18	362.5	78.3	105	1	INS_SHEEP	P01318 ovies arie
19	334.5	72.2	108	1	INS2_COWP	P21553 capra hirc
20	314	67.4	110	1	INS_CAVPO	P01329 capra hirc
21	277.5	59.5	105	1	INS_COTR	P07715 actinon ocy
22	276.5	59.2	103	1	INS2_BELF	P01414 belapadus
23	265.5	57.3	106	1	INS2_XENLA	P12732 xenopus lae
24	265.5	57.3	107	1	INS2_CHICK	P01272 gallus gal
25	262.5	56.7	106	1	INS1_XENLA	P12736 xenopus lae
26	256.5	55.4	51	1	INS2_BALPH	P01312 balancerp
27	256.5	55.4	51	1	INS_ELEWA	P01316 elephas max
28	256.5	55.4	81	1	INS2_WAPD	P01333 anas platyr
29	251.5	54.3	51	1	INS2_ACOCA	P01324 acorys cal
30	246.5	53.2	51	1	INS2_BALBO	P01314 balaeopte
31	246.5	53.2	51	1	INS2_CAMDR	P01320 camelus d
32	246.5	53.2	51	1	INS2_CAPRI	P01319 capra hirc
33	245.5	53.0	51	1	INS2_FELCA	P06306 felis silve

44	239.5	51.2	51	1	INS_LIZMA	P18109 didelphis m
35	234.5	50.6	51	1	INS_CHIBR	P01327 chinchilla
36	231.5	50.3	51	1	INS_ARSAN	P07454 anser anser
37	228	49.2	118	1	INS_VERMO	O96752 verasper mo
38	227.5	49.1	51	1	INS_HVSCR	P01328 hystrix cri
39	227.5	49.1	51	1	INS_TRASC	P31857 trachemys s
40	225	48.6	52	1	INS_ACIGU	P01423 acipenser g
41	221.5	47.8	116	1	INS_LOPPI	P01341 lophius pis
42	218.5	47.2	51	1	INS_ORNAN	O96757 ornithorhyn
43	218.5	47.2	105	1	INS_ONCKE	P04667 oncorhynch
44	216.5	46.8	89	1	INS_CALMI	P13190 callorhynch
45	215	46.4	113	1	INS_OREN3	P01025 oreochromis

ALIGNMENTS

RESULT 1	INS_HUMAN	STANDARD	PRT: 110 AA.
AC	P01309:		
DT	21-JUL-1986 (Rel. 02, Created)		
DT	21-JUL-1986 (Rel. 01, Last sequence update)		
DT	15-SEP-2003 (Rel. 42, Last annotation update)		
CE	Insulin precursor.		
GN	INS.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.		
OX	NCBI:taxid=9606;		
RP	SEQUENCE FROM N.A.		
RP	MEDLINE=80120725; PubMed=6243748;		
RA	Bell G.I., Pictet R.L., Rutter W.J., Cordell B., Tischer E.,		
RA	Goodman H.M.;		
RL	"Sequence of the human insulin gene."		
RL	Nature 284:26-32(1985).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RP	MEDLINE=80236133; PubMed=6248962;		
RA	Ulrich A., Dull T.J., Gray A., Brosius J., Sures I.,		
RA	"Genetic variation in the human insulin gene."		
RA	Science 209:612-615(1983).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RP	MEDLINE=80054779; PubMed=503234;		
RA	Bell G.I., Swain W.F., Pictet R.L., Cordell B., Goodman H.M.,		
RA	Rutter W.J.;		
RL	"Nucleotide sequence of a cDNA clone encoding human preproinsulin."		
RL	Nature 262:525-527(1979).		
RN	[4]		
RP	SEQUENCE FROM N.A.		
RP	MEDLINE=80147417; PubMed=6227840;		
RA	Soares I., Goeddel D.V., Gray A., Ulrich A.;		
RA	"Nucleotide sequence of human preproinsulin complementary DNA."		
RA	Science 208:57-59(1980).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RP	MEDLINE=83364428; PubMed=8358440;		
RA	Leussen A.M., Bell T., Julier C., Latour M.;		
RA	"Susceptibility to insulin dependent diabetes mellitus maps to a 4.1		
PT	kb segment of DNA spanning the insulin gene and associated VNTR."		
PT	Nat. Genet. 4:305-310(1993).		
RN	[6]		
RP	SEQUENCE FROM N.A.		
RP	TISSUE=pancreas;		
RP	MEDLINE=22388257; PubMed=12477932;		
RA	Straussberg R., Feingold E.A., Grouse L.H., Berge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,		
RA	Altshuler S.F., Zeeberg K.H., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,		
RA	Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,		
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,		

RA Brownstein M.J., Usdin T.B., Teshiyuki S., Carninci P., Prange G.,
 RA Raha S.S., Loquerano N.A., Peters G.J., Abramson R.D., Mellichy S.C.,
 RA Bosak S.A., McSwan P.J., McEwen K.J., Malek J.A., Guetart P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay J., Hulyk S.W.,
 RA Villalon D.K., Muzay D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Ketteran M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard J.G.,
 RA Blakesley R.W., Touchman J.W., Grein E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers K.M.,
 RA Butterfield V.S.N., Krzywinski M.J., Skalska H., Smailus D.E.,
 RA Schnerch A., Schell J.E., Jones S.J.M., Karra R.A.,
 RA "Generation and initial analysis of more than 55,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16669-16693(2002).
 RN [7]
 RN SEQUENCE OF 1-59 FROM N.A.
 RC TISSUE=Blood;
 RA Fajardy I., Weill J.J., Stuckens C.C., Danze P.M.F.;
 RA "Description of a novel RFLP diallelic polymorphism (-127 HpaI G/G)
 RT within the 5' region of insulin gene.";
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RN SEQUENCE OF 25-54 AND 90-110.
 RA Nicol D.S.H., Smith I.F.;
 RA "Amino-acid sequence of human insulin.";
 RT Nature 187:483-485(1966).
 RN [9]
 RN SEQUENCE OF 57-87.
 RA MEDLINE=71116410; PubMed=5101771;
 RA Oyer E.E., Che S., Petersen D.B., Steiner D.B.;
 RA "Studies on human proinsulin: Isolation and amino acid sequence of
 RT the human pancreatic C-peptide.";
 RL J. Biol. Chem. 246:1375-1386(1971).
 RN [10]
 RN SEQUENCE OF 57-87.
 RA MEDLINE=71257722; PubMed=5563464;
 RA KC A., Smyth D.G., Markussen J., Sindhya P.;
 RA "The amino acid sequence of the C-peptide of human proinsulin.";
 RT Eur. J. Biochem. 20:199-199(1971).
 RN [11]
 RN SYNTHESIS.
 RA MEDLINE=75077277; PubMed 4444293;
 RA Steber P., Kamber B., Hartmann A., Jeoll A., Rin, K., Ritzel W.,
 RA "Total synthesis of human insulin under directed formation of the
 RT disulfide bonds.";
 RL Helv. Chim. Acta 57:2617-2621(1974).
 RN [12]
 RN SYNTHESIS OF 57-87.
 RA MEDLINE=75040007; PubMed=4833504;
 RA Naithani V.K.;
 RA "Studies on polypeptides, IV. The synthesis of C-peptide of human
 RT proinsulin.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 354:659-672(1973).
 RN [13]
 RN SYNTHESIS OF 65-69 AND 70-73.
 RA MEDLINE=73161263; PubMed=4698555;
 RA Geiger R., Volk A.;
 RA "Synthesis of peptides with the properties of human proinsulin C
 RT peptides (hc peptide). 3. Synthesis of the sequences 14-17 and 6-13
 RI of human proinsulin C peptides.";
 RL Chem. Ber. 106:199-205(1973).
 RN [14]
 RN SYNTHESIS OF 84-87.
 RA MEDLINE=73161261; PubMed=4698553;
 RA Geiger R., Geiger G., Keck W., Treuth G.;
 RA "Synthesis of peptides with the properties of human proinsulin C
 RT peptides (hc peptide). 1. Scheme for the synthesis and preparation of
 RI the sequence 28-31 of human proinsulin C peptide.";
 RL Chem. Ber. 106:186-192(1973).
 RN [15]
 RN VARIANT: LOS ANGELES SER-48.
 RA MEDLINE=84016053; PubMed=6312459;
 RA Haneda M., Chan S.J., Kwok S.C.M., Rubenstein A.H., Steiner D.F.;
 RA "Studies on mutant human insulin genes: identification and sequence
 RT analysis of a gene encoding [Ser24]insulin.";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:6366-6370(1983).
 RN [16]
 RN VARIANTS LOS ANGELES SER-46 AND CHICAGO LEU-49.
 RA MEDLINE=84170233; PubMed=6424111;
 RA Shoenelson S., Fickova M., Haneda M., Nahum A., Musso G., Kaiser E.T.,
 RA Rubenstein A.H., Tager H.;
 RA "Identification of a mutant human insulin: predicted to contain a
 RT serine-for-phenylalanine substitution.";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:7390-7394(1983).
 RN [17]
 RN VARIANT PROVIDENCE ASP-34.
 RA MEDLINE=8775640; PubMed=3470784;
 RA Chan S.J., Seino S., Gruppuso P.A., Schwartz R., Steiner D.F.;
 RA "A mutation in the B chain coding region is associated with impaired
 RT proinsulin conversion in a family with hyperproinsulinemia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:2194-2197(1987).
 RN [18]
 RN VARIANT WAKAYAMA LEU-92.
 RA MEDLINE=87058122; PubMed=3537011;
 RA Sakura H., Iwamoto Y., Sakamoto Y., Kuzuya T., Hirata H.;
 RA "Structurally abnormal insulin in a diabetic patient. Characterization
 RT of the mutant insulin A3 (Val->Leu) isolated from the pancreas.";
 RL J. Clin. Invest. 78:1668-1672(1986).
 RN [19]
 RN VARIANT HIS-89.
 RA MEDLINE=90317021; PubMed=2196379;
 RA Barbetti F., Raben N., Kadosaki T., Cama A., Accili D., Gabbay K.H.,
 RA Meyerich J.A., Taylor S.J., Roth J.;
 RA "Two unrelated patients with familial hyperproinsulinemia due to a
 RT mutation substituting histidine for arginine at position 65 in the
 RI proinsulin molecule: identification of the mutation by direct
 R2 sequencing of genomic deoxyribonucleic acid amplified by polymerase
 R3 chain reaction.";
 RL J. Clin. Endocrinol. Metab. 71:164-169(1990).
 RN [20]
 RN VARIANT HIS-89.
 RA MEDLINE=85261996; PubMed=4019786;
 RA Shibasaki Y., Kawakami T., Kanazawa Y., Akanuma Y., Takaku F.;
 RA "Posttranslational cleavage of proinsulin is blocked by a point
 RT mutation in familial hyperproinsulinemia.";
 RL J. Clin. Invest. 76:374-380(1985).
 RN [21]
 RN VARIANT KYOTO LEU-89.
 RA MEDLINE=92291507; PubMed=1601597;
 RA Yano H., Kitano N., Morimoto M., Polonsky K.S., Izura H., Seino Y.;
 RA "A novel point mutation in the human insulin gene giving rise to
 RT hyperproinsulinemia (proinsulin Kyoto).";
 RL J. Clin. Invest. 89:1902-1907(1992).
 RN [22]
 RN STRUCTURE BY NMR.
 RA MEDLINE=91104956; PubMed=2271664;
 RA Hua Q.-X., Weiss M.A.;
 RA "Toward the solution structure of human insulin: sequential 2D 1H NMR
 RT assignment of a des-pentapeptide analogue and comparison with crystal
 RL structure.";
 RL Biochemistry 29:10545-10555(1990).
 RN [23]
 RN STRUCTURE BY NMR.
 RA MEDLINE=91242467; PubMed=2036420;
 RA Hua Q.-X., Weiss M.A.;
 RA "Comparative 2D NMR studies of human insulin and des-pentapeptide
 RT insulin: sequential resonance assignment and implications for protein
 RI dynamics and receptor recognition.";
 RL Biochemistry 30:5505-5515(1991).
 RN [24]
 RN STRUCTURE BY NMR.
 RA MEDLINE=91265527; PubMed=1546635;
 RA Hua Q.-X., Weiss M.A.;
 RA "Two-dimensional NMR studies of Des-(B26-B30)-insulin: sequence-
 RT specific resonance assignments and effects of solvent composition.";
 RL Biochim. Biophys. Acta 1078:101-110(1991).

```

Query Match      100.0%; Score 453; 25 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 9.2e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Query Match      98.5%   Score 456;   DB 1;   Length 110;
Best Local Similarity 98.8%   Pred. No. 5,2e-42;
Matches 85;   Conservative 0;   Mismatches 1;   Indels 6;   Gaps 0;

1 FVNOHCGSHVEALYLVCGGGEFFYTPKTRREAEDLVGVGEGGPGAGSLOPLALEG 60
:|||||1.|||||1.|||||1.|||||1.|||||1.|||||1.|||||1.|||||1.
25 FVNOHCGSHVEALYLVCGGGEFFYTPKTRREAEDLVGVGEGGPGAGSLOPLALEG 84
:|||||1.|||||1.|||||1.|||||1.|||||1.|||||1.|||||1.|||||1.

QY 61 SLQKRGIVECCCTSCISLYQLENYCN 86
:|||||1.|||||1.|||||1.|||||1.|||||1.|||||1.|||||1.
DB 85 SLQKRGIVECCCTSCISLYQLENYCN 110
:|||||1.|||||1.|||||1.|||||1.|||||1.|||||1.|||||1.

RESULT 5
INS_MACFA
ID INS_MACFA STANDARD: PRT: 110 AA.
AC P30406; P01303;
DT 21-JUL-1986 (Rel. 01; Created)
DT 13-AUG-1987 (Rel. 05; Last sequence update)
DT 01-FEB-1996 (Rel. 33; Last annotation update)
DE Insulin precursor.
GN INS.
OS Macaca fascicularis (Cran eating macaque) (Cynemolagus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8308074; PubMed=6164262;
RA Wellekam W., Groeber J., Jelinek M., Willemsen F.,
RA Winnacker E.-L.;
RT The nucleotide sequence of cDNA coding for preproinsulin from the
RT primate Macaca fascicularis.
RL Gene 13:119-183(1982).
CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J00336; AAA16849.1;
DR PIR: J00178; JC0178.
DR HSP: P01308; IAI0.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00649; Insulin.
DR PRINTS: PR00277; INSULIN.
DR SMART: SM00078; IIGF.
DR PROSITE: PS00262; INSULIN.
KW Insulin family; Hormone; Glucose metabolism; Signal.
FT SIGNAL 1 24
FT CHAIN 25 54 INSULIN B CHAIN.
FT PROPEP 57 87 C PEPTIDE.
FT CHAIN 90 110 INSULIN A CHAIN.
FT DISULFD 31 96 INTERCHAIN.
FT DISULFD 43 103 INTERCHAIN.
FT DISULFD 95 100
SQ SEQUENCE 110 AA; 11991 MW; 83C6E13A63A420E9 CRC64;

Query Match      98.5%   Score 456;   DB 1;   Length 110;
Best Local Similarity 98.8%   Pred. No. 5,2e-42;
Matches 85;   Conservative 0;   Mismatches 1;   Indels 6;   Gaps 0;

1 FVNOHCGSHVEALYLVCGGGEFFYTPKTRREAEDLVGVGEGGPGAGSLOPLALEG 60
:|||||1.|||||1.|||||1.|||||1.|||||1.|||||1.|||||1.
25 FVNOHCGSHVEALYLVCGGGEFFYTPKTRREAEDLVGVGEGGPGAGSLOPLALEG 84
:|||||1.|||||1.|||||1.|||||1.|||||1.|||||1.|||||1.

QY 61 SLQKRGIVECCCTSCISLYQLENYCN 86
:|||||1.|||||1.|||||1.|||||1.|||||1.|||||1.|||||1.
DB 85 SLQKRGIVECCCTSCISLYQLENYCN 110
:|||||1.|||||1.|||||1.|||||1.|||||1.|||||1.|||||1.

RESULT 5
INS_MACFA
ID INS_MACFA STANDARD: PRT: 110 AA.
AC P30406; P01303;
DT 21-JUL-1986 (Rel. 01; Created)
DT 13-AUG-1987 (Rel. 05; Last sequence update)
DT 01-FEB-1996 (Rel. 33; Last annotation update)
DE Insulin precursor.
GN INS.
OS Macaca fascicularis (Cran eating macaque) (Cynemolagus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8308074; PubMed=6164262;
RA Wellekam W., Groeber J., Jelinek M., Willemsen F.,
RA Winnacker E.-L.;
RT The nucleotide sequence of cDNA coding for preproinsulin from the
RT primate Macaca fascicularis.
RL Gene 13:119-183(1982).
CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J00336; AAA16849.1;
DR PIR: J00178; JC0178.
DR HSP: P01308; IAI0.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00649; Insulin.
DR PRINTS: PR00277; INSULIN.
DR SMART: SM00078; IIGF.
DR PROSITE: PS00262; INSULIN.
KW Insulin family; Hormone; Glucose metabolism; Signal.
FT SIGNAL 1 24
FT CHAIN 25 54 INSULIN B CHAIN.
FT PROPEP 57 87 C PEPTIDE.
FT CHAIN 90 110 INSULIN A CHAIN.
FT DISULFD 31 96 INTERCHAIN.
FT DISULFD 43 103 INTERCHAIN.
FT DISULFD 95 100
SQ SEQUENCE 110 AA; 11991 MW; 83C6E13A63A420E9 CRC64;

```

```

FT DISULFID 31 96 INTERCHAIN.
FT DISULFID 43 109 INTERCHAIN.
FT DISULFID 95 100
FT CONFLICT 83 85 E - - Y (IN REF. 8)
SQ SEQUENCE 110 AA: 11838 MW: 82029/53855/7788 CRC64:

Query Match
Best Local Similarity 90.74; Score 424; DB 1; Length 110;
Matches 78; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVEAFLYVGCGHGFYTPKTRREAEALQVQVLEGGGAGSLQPLALEG 60
Db 25 FVNHLCGSHLVEAFLYVGCGHGFYTPKTRREAEALQVQVLEGGGAGSLQPLALEG 64

QY 61 SLQKRGIVGQCTSTCSLYQLENYCN 86
Db 85 ALQKRGIVGQCTSTCSLYQLENYCN 110

RESULT 6
INS_CANFA
ID INS_CANFA STANDARD: PRT: 110 AA.
AC P01321:
DT 21-JUL-1986 (Rel. 61, Created)
DI 21-JUL-1986 (Rel. 61, Last sequence update)
DE 01-FEB-1996 (Rel. 33, Last annotation update)
GN INS
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83109071; PubMed=6296142;
RA Kwok S.C.M., Chan S.J., Steiner D.P.;
RT "Cloning and nucleotide sequence analysis of the dog insulin gene.
RT Coded amino acid sequence of canine preproinsulin products an
RT additional C-peptide fragment."
RL J. Biol. Chem. 258:2357-2364(1983).
RN [2]
RP SEQUENCE OF 25 54 AND 90-110.
RX MEDLINE=66160119; PubMed=5949503;
RA Smith L.F.;
RT "Species variation in the amino acid sequence of insulin."
RL Am. J. Med. 40:662-666(1966).
CC -1- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: V00179; CAA23475.1;
CC PIR: A92413; IPDG.
CC HSSP: P01317; IAPG.
CC InterPro: IPR004325; Ins/IGF/relax.
CC Pfam: PF00649; Insulin; 1.
CC SMART: SM00078; IIGF; 1.
CC PROSITE: PS00262; INSULIN; 1.
CC Insulin family: Hormone; Glucose metabolism; Signal.
CC SIGNAL 1 24
CC CHAIN 25 54 INSULIN B CHAIN.
CC
CC DR PIR: A92413; IPDG.
CC DR HSSP: P01317; IAPG.
CC DR InterPro: IPR004325; Ins/IGF/relax.
CC DR Pfam: PF00649; Insulin; 1.
CC DR SMART: SM00078; IIGF; 1.
CC DR PROSITE: PS00262; INSULIN; 1.
CC Insulin family: Hormone; Glucose metabolism; Signal.
CC SIGNAL 1 24
CC CHAIN 25 54 INSULIN B CHAIN.

```

```

FT PROPEP 57 87 C PEPTIDE.
FT CHAIN 90 110 INSULIN A CHAIN.
FT DISULFID 31 96 INTERCHAIN.
FT DISULFID 43 109 INTERCHAIN.
FT DISULFID 95 100
SQ SEQUENCE 110 AA: 12190 MW: 45747918644/4F98 CRC64:

Query Match
Best Local Similarity 89.58; Score 417; DB 1; Length 110;
Matches 77; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVEAFLYVGCGHGFYTPKTRREAEALQVQVLEGGGAGSLQPLALEG 60
Db 25 FVNHLCGSHLVEAFLYVGCGHGFYTPKTRREAEALQVQVLEGGGAGSLQPLALEG 84

QY 61 SLQKRGIVGQCTSTCSLYQLENYCN 85
Db 85 ALQKRGIVGQCTSTCSLYQLENYCN 110

RESULT 7
INS_SPETR
ID INS_SPETR STANDARD: PRT: 110 AA.
AC Q91X13:
DI 15-SEP-2003 (Rel. 42, Created)
DI 15-SEP-2003 (Rel. 42, Last sequence update)
DI 15-SEP-2003 (Rel. 42, Last annotation update)
DE Insulin precursor.
GN INS.
OS Spromophilus tridecemlineatus (Thirteen-lined ground squirrel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;
OC Spromophilus.
OX NCBI_TaxID=43179;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83109071; PubMed=6296142;
RA Kwok S.C.M., Chan S.J., Steiner D.P.;
RT "Cloning and nucleotide sequence analysis of the dog insulin gene.
RT Coded amino acid sequence of canine preproinsulin products an
RT additional C-peptide fragment."
RL J. Biol. Chem. 258:2357-2364(1983).
RN [2]
RP SEQUENCE OF 25 54 AND 90-110.
RX MEDLINE=66160119; PubMed=5949503;
RA Smith L.F.;
RT "Species variation in the amino acid sequence of insulin."
RL Am. J. Med. 40:662-666(1966).
CC -1- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AY038604; AAK72558.1;
CC HSSP: P01308; IAPG.
CC InterPro: IPR004325; Ins/IGF/relax.
CC Pfam: PF00649; Insulin; 1.
CC SMART: SM00078; IIGF; 1.
CC PROSITE: PS00262; INSULIN; 1.
CC Insulin family: Hormone; Glucose metabolism; Signal.
CC SIGNAL 1 24
CC CHAIN 25 54 INSULIN B CHAIN.
CC
CC DR PIR: A92413; IPDG.
CC DR HSSP: P01317; IAPG.
CC DR InterPro: IPR004325; Ins/IGF/relax.
CC DR Pfam: PF00649; Insulin; 1.
CC DR SMART: SM00078; IIGF; 1.
CC DR PROSITE: PS00262; INSULIN; 1.
CC Insulin family: Hormone; Glucose metabolism; Signal.
CC SIGNAL 1 24
CC CHAIN 25 54 INSULIN B CHAIN.

```

Query Match 89.24; Score 423; [R 1; Length 110;

Best Local Similarity 89.59; Pred. No. 20-37;

Matches 77; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRRAEDQLVGVVQLGGPFGAGSLQPLALEG 61

DB 25 FVNHLCGSHLVEALYLVCGERGFFYTPKSRREVERGCGGVELGCGFGAGLQPLALEM 84

QY 61 SLQKRGIVECCSTCSLYLENYCN 86

DB 85 ALQKRGIVECCSTCSLYLENYCN 110

RESULT 8

INS_HORSE STANDARD: PRT: 86 AA.

AC PG1310;

DT 21-JUL-1986 (Rel. 31, Created)

DT 21-JUL-1986 (Rel. 31, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Insulin precursor.

GN INS.

OS Equus caballus (Horse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

OX NCBI_TaxID:9796;

RN (1)

RP SEQUENCE OF 1-30 AND 66-85

RA Harris J., Santer P., Naughton M.A.;

RT "Species differences in insulin."

RL Arch. Biochem. Biophys. 65:427-438(1989).

RN (2)

RP SEQUENCE OF 33-65.

RX MEDLINE-73061498; PubMed-4649931;

RA Tager H.S., Steiner D.F.;

RT "Primary structures of the proinsulin connecting peptides of the rat and the horse."

RL J. Biol. Chem. 247:7936-7940(1972).

CC -1- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT

CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND

CC FATTY ACIDS. IT ACCELERATES BLOOD GLUCOSE CONCENTRATION. IT

CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.

CC -1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO

CC DISULFIDE BONDS.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.

CC -1- CAUTION: X'S AT POSITIONS 31-32 AND 64-65 REPRESENT PATTERNED BASIC

CC RESIDUES ASSUMED BY HOMOLGY TO BE PRESENT IN THE PRECURSOR

CC MOLECULE.

PIR: AC1590; IAPH.

DR HSSP; P01317; IAPH.

DR InterPro: IPR004825; Ins/IGF/relax.

DR Pfam: PF00049; Insulin; 1.

DR SMART: SM00078; ILGF; 1.

DR PROSITE: PS00262; INSULIN; 1.

KW Insulin family; Hormone; Glucose metabolism.

FT CHAIN 1 30 INSULIN B CHAIN.

FT PROPEP 33 63 C PEPTIDE.

FT CHAIN 66 85 INSULIN A CHAIN.

FT DISULFID 7 72 INTERCHAIN.

FT DISULFID 19 85 INTERCHAIN.

FT DISULFID 71 76

SQ SEQUENCE 86 AA; 9142 MW; A3E1E822711B0546 CRC64;

Query Match

Best Local Similarity 85.18; Score 394; [R 1; Length 86;

Matches 73; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRRAEDQLVGVVQLGGPFGAGSLQPLALEG 61

DB 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRRAEDQLVGVVQLGGPFGAGSLQPLALEG 61

QY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRRAEDQLVGVVQLGGPFGAGSLQPLALEG 61

DB 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRRAEDQLVGVVQLGGPFGAGSLQPLALEG 61

QY 61 SLQKRGIVECCSTCSLYLENYCN 86

DB 61 PQXXGIVECCSTCSLYLENYCN 86

RESULT 9

INS2_MOUSE STANDARD: PRT: 110 AA.

ID INS2_MOUSE

AC P01326;

DT 21-JUL-1986 (Rel. 31, Created)

DT 13-AUG-1987 (Rel. 05, Last sequence update)

DT 15-OCT-2001 (Rel. 40, Last annotation update)

DE Insulin 2 precursor.

GN INS2 OR INS-2

OS Mus musculus (Mouse)

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI_TaxID:10090;

RN (1)

RP SEQUENCE FROM N.A.

RX MEDLINE-67169758; PubMed-3104603;

RA Wentworth B.M., Schaefer I.M., Villa-Komaroff L., Chirgwin J.M.;

RT "Characterization of the two nonallelic genes encoding mouse

RT Preproinsulin."

PL J. Mol. Evol. 23:305-312(1985).

RN (2)

RP SEQUENCE FROM N.A.

RX STRAIN-MON.

RX MEDLINE-90372959; PubMed-2397023;

RA Sawa T., Ogaku S., Moriocka H., Yano S.;

RT "Molecular cloning and DNA sequence analysis of preproinsulin genes

RT in the NON mouse, an animal model of human non-obese, non-insulin-

RT dependent diabetes mellitus."

RL J. Mol. Endocrinol. 5:61-67(1990).

RN (3)

RP SEQUENCE OF 25-54 AND 90-110.

RX MEDLINE-72189455; PubMed-5063715;

RA Buentzli H.F., Glatthar B., Kunz P., Muehlaupt E., Hummel R.E.;

RT "Amino acid sequence of the two insulins from mouse (Maus musculus)."

RL Hoppe-Seivler's Z. Physiol. Chem. 353:451-458(1972)

CC -1- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT

CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND

CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE

CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.

CC -1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO

CC DISULFIDE BONDS.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.

CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC EMBL: X04724; CAA28433.1; -

DR PIR: A26342; INMS2

DR HSSP: P01317; IAPH.

DR KSD; MGI:96573; Ins2.

DR GO: GO:0000187; P:activation of MAPK; IDA.

DR GO: GO:0042325; P:regulation of phosphorylation; IDA.

DR InterPro: IPR004825; Ins/IGF/relax.

DR Pfam: PF00049; Insulin; 1.

DR SMART: SM00078; ILGF; 1.

DR PROSITE: PS00262; INSULIN; 1.

KW Insulin family; Hormone; Glucose metabolism; Signal; Multigene family.

FT CHAIN 1 24

FT CHAIN 25 54 INSULIN 2 B CHAIN.

FT PROPEP 57 87 INSULIN 2 C PEPTIDE.

FT CHAIN 90 110 INSULIN 2 A CHAIN.

FT DISULFID 31 96 INTERCHAIN.


```

FT DISULFID 43 105 INTERCHAIN.
FT DISULFID 95 100
SQ SEQUENCE 110 AA: 12364 MA: 3564CR80:DE24:PLAG CR804.

Query Match
Best Local Similarity 85.1%; Score 394; DB 1; Length 110;
Matches 73; Conservative 4; Mismatches 9; Indels 0; Gaps 0.

CY 1 FVNQHGHSHLVEALYLVCGERFYYTKPREAEADLVAGQVELGGPGAGSLQPLALEG 60
DB 25 FVKHLCGSHLVEALYLVCGERFYYTKPREAEADLVAGQVELGGPGAGSLQPLALEG 60
CY 61 SLCKKGIVGQVCTSCSLYLENYCN 86
DB 35 AOKKRGIVGQVCTSCSLYLENYCN 110

RESULT 10
INS2_RAT
ID INS2_RAT STANDARD: PRT: 110 AA.
AC P01323;
DI 21-JUL-1986 (Rel. 01, Created)
DI 21-JUL-1986 (Rel. 01, Last sequence update)
DI 16-OCT-2001 (Rel. 40, Last annotation update)
DE Insulin; 2 precursor.
GN INS2 OR INS-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-Sprague-Dawley; TISSUE=Liver;
RC MEDLINE=80045035; PubMed=496284;
RA Lomedico P., Rosenthal N., Bistratiadis A., Gilbert W., Kolodner R.,
RA Tizard R.;
R1 "The structure and evolution of the two nonallelic rat preproinsulin
R1 genes.";
RL Cell 18:545-558(1979).
[2]
RN SEQUENCE FROM N.A.
RP MEDLINE=86310882; PubMed=2427932;
RA Soares M.B., Schin E., Henderson A., Karathanasis S.K., Gale R.,
RA Zeitlin S., Charwin S., Bistratiadis A.;
R1 "RNA-mediated gene duplication: the rat preproinsulin 1 gene is a
R1 functional retroposon.";
RL Mol. Cell. Biol. 5:2096-2103(1985).
[3]
RN SEQUENCE FROM N.A.
RP MEDLINE=80240375; PubMed=6743167;
RA Lomedico P., Rosenthal N., Kolodner R., Bistratiadis A.,
RA Gilbert W.;
R1 "The structure of rat preproinsulin genes.";
RL Acc. N.Y. Acad. Sci. 343:425-432(1980).
[4]
RN SEQUENCE OF 25-54 AND 90-110.
RP MEDLINE=70067613; PubMed=4311948;
RA Steiner D.F., Clark J.L., Nolan C., Roberstein A.D., Margellash E.,
RA Aten B., Over P.E.;
R1 "Proinsulin and the biosynthesis of insulin.";
RL Recent Prog. Horm. Res. 25:207-282(1969).
[5]
RN SEQUENCE OF 57-87.
RP MEDLINE=73061498; PubMed=4640931;
RA Tager H.S., Steiner D.F.;
R1 "Primary structures of the proinsulin connecting peptides of the rat
R1 and the horse.";
RL J. Biol. Chem. 247:7936-7940(1972).
[6]
RN SEQUENCE OF 57-87, AND REVISED.
RP MEDLINE=72:77385; PubMed=4554104;
RA Markussen J., Sundby F.;
R1 "Rat-proinsulin C-peptides. Amino-acid sequences.";

```

```

Sur. J. Biochem. 25:153-162(1972).
-1- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
-1- SUBUNIT: HETEROOLIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
DISULFIDE BONDS.
-1- SUBCELLULAR LOCATION: Secreted.
-1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL: V01243; CAA24560.1;
EMBL: J00748; AAA41443.1;
EMBL: M25585; AAA41440.1;
EMBL: M25583; AAA41440.1; JOINED.
PIR: B90789; IPRT2.
HSP: P01317; IAPH.
InterPro: IPR004825; Ins/IGF/relax.
Pfam: PF00049; Insulin; 1.
SMART: SM00078; IIGF; 1.
PROSITE: PS00262; INSULIN; 1.
Insulin family; Hormone; Glucose metabolism; Signal; Multigene family.
FT SIGNAL 1 24
FT CHAIN 25 54 INSULIN 2 B CHAIN.
FT PROPEP 57 87 INSULIN 2 C PEPTIDE.
FT CHAIN 90 110 INSULIN 2 A CHAIN.
FT DISULFID 31 96 INTERCHAIN.
FT DISULFID 43 109 INTERCHAIN.
FT DISULFID 95 100
SQ SEQUENCE 110 AA: 12339 MW: 3A626DA98C86F3CA CRC64:

Query Match
Best Local Similarity 85.1%; Score 394; DB 1; Length 110;
Matches 73; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

CY 1 FVNQHGHSHLVEALYLVCGERFYYTKPREAEADLVAGQVELGGPGAGSLQPLALEG 60
DB 25 FVKHLCGSHLVEALYLVCGERFYYTKPREAEADLVAGQVELGGPGAGSLQPLALEG 64
CY 61 SLCKKGIVGQVCTSCSLYLENYCN 86
DB 35 AOKKRGIVGQVCTSCSLYLENYCN 110

RESULT 11
INS_RAT
ID INS_RAT STANDARD: PRT: 108 AA.
AC P10604;
DI 01-JUL-1989 (Rel. 11, Created)
DI 01-JUL-1989 (Rel. 11, Last sequence update)
DI 16-OCT-2001 (Rel. 40, Last annotation update)
DE Insulin precursor.
GN INS.
OS Aotus trivirgatus (Night monkey) (Deomouculi), and
OS Aotus trivirgatus (Common squirrel monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Aotinae; Aotus.
OX NCBI_TaxID=9505, 9521;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=A. trivirgatus;
RX MEDLINE=88041119; PubMed=3118367;
RA Seino S., Steiner D.F., Bell G.I.;
R1 "Sequence of a New World primate insulin having low biological
R1 potency and immunoreactivity.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:7423-7427(1987).

```


Cell 18:523-542(1979).
 [2] SEQUENCE FROM N.A.
 RC STRAIN-SQUAMOUS-Dawley; T18-DB-11907;
 RA MEDLINE=60949033; PubMed=454284;
 RA Lomedico P., Rosenthal N., Eistatidis A., Gilbert W., Kothmann P.,
 RA Tizard R.;
 KT "The structure and evolution of the two homologous rat preproinsulin
 genes";
 RL Cell 18:545-558(1979).
 RL [3]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=80240379; PubMed=6249167;
 RA Lomedico P.L., Rosenthal N., Kothmann P., Eistatidis A.,
 RA Gilbert W.;
 RT "The structure of rat preproinsulin genes";
 RL Ann. N.Y. Acad. Sci. 343:425-432(1980).
 RN [4]
 RN SEQUENCE OF 25-54 AND 90-110.
 RP MEDLINE=70067613; PubMed=4111936;
 RA Steiner D.F., Clark J.L., Nolan C., Rubenstein A.H., Matsuzaki E.,
 RA Aten B., Oyer P.E.;
 RT "Proinsulin and the biosynthesis of insulin";
 RL Recent Prog. Horm. Res. 25:207-282(1969).
 RN [5]
 RP SEQUENCE OF 57-87.
 RX MEDLINE=73061499; PubMed=4640931;
 RA Tager H.S., Steiner D.F.;
 RT "Primary structures of the proinsulin connecting peptides of the rat
 and the horse";
 RL J. Biol. Chem. 247:7935-7940(1972).
 RN [6]
 RP SEQUENCE OF 57-87, AND REV-STONS.
 RX MEDLINE=72177395; PubMed=4554104;
 RA Markussen J., Sundby F.;
 RT "Rat-proinsulin C-peptides. Amino-acid sequences.";
 RL Eur. J. Biochem. 25:153-162(1972).
 CC -1- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
 CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
 CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
 CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
 CC -1- SUBUNIT: HETERODIMER OF A H CHAIN AND AN A CHAIN LINKED BY TWO
 CC DISULFIDE BONDS.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE "INSULIN/IGF/RELAXIN FAMILY".
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.sib.ch/usage/>
 CC or send an email to license@isb.sib.ch).
 CC
 DR EMBL: V01242; CAA24559.1;
 DR EMBL: J00747; AAA41442.1;
 DR EMBL: M25584; AAA41439.1;
 DR PIR: A30786; IPRT1.
 DR HSSP: PC1308; 1A7F.
 DR InterPro: IPR0034825; Ins/IGF/relax.
 DR Pfam: PF00049; Insulin; 1.
 DR SMART: SM00078; IIGF; 1.
 DR PROSITE: PS02624; INSULIN; 1.
 KW Insulin family; Hormone; Glucose metabolism; Signal; Multigene family.
 FT SIGNAL 1 24
 FT CHAIN 25 54 INSULIN 1 B CHAIN.
 FT PROPEP 57 87 INSULIN 1 C PEPTIDE.
 FT CHAIN 90 110 INSULIN 1 A CHAIN.
 FT DISULFID 31 96 INTERCHAIN.
 FT DISULFID 43 109 INTERCHAIN.
 FT DISULFID 95 100
 SQ SEQUENCE 110 AA: 12420 WK: 51D606DA54AE4533 CRC64:

Query Match 63.2% Score 385; DB 1; Length 110;
 Best Local Similarity 83.7% Pred. No. 2e-34;
 Matches 72; Conservative 4; Mismatches 10; Gaps 0;
 0Y 1 FYNHLCGSHAEVLYVCGERGFFYTPKTRAEEDLVGVVELGGGFCASSLQPLALEG 60
 11
 25 FVQKHLCGCPHEVLYVCGERGFFYTPKSRREVEDQVQPLELGGGPEAGPLQTLALPV 84
 0Y 61 SLOKSGIVEQCCTSCSLYQLENYCN 86
 11
 85 ARKRGIVQCCTSCSLYQLENYCN 110
 RESULT 14
 INS_P56 STANDARD; PRT: 108 AA.
 CC INS_P13
 AC 50:315; 091535;
 DC 21-JUL-1986 (Rel. 02, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Insulin precursor.
 DE INS.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Han X.G., Tsch S.E.;
 RT "Complete porcine preproinsulin cDNA sequence.";
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBS3 databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Large white;
 RX MEDLINE=22135958; PubMed=12140686;
 RA Amarger V., Nguyen M., Laere A.S., Braunschweig M., Nezer C.,
 RA Georges M., Andersson L.;
 RT "Comparative sequence analysis of the INS-IGF2-H19 gene cluster in
 RT pigs.";
 RL Mamm. Genome 13:388-398(2002).
 RN [3]
 RP SEQUENCE OF 25-108.
 RX MEDLINE=6826485; PubMed=5657063;
 RA Chance R.E., Ellis K.M., Bromer W.W.;
 RT "Porcine proinsulin: characterization and amino acid sequence.";
 RL Science 161:165-167(1968).
 RN [4]
 RP REVISION TO 59.
 RA Chance R.E.;
 RL Submitted (JUL-1970) to the PIR data bank.
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
 RA Blundell T.L., Dodson G.G., Hodgkin D., Mercola D.;
 RT "Insulin: the structure in the crystal and its reflection in
 RT chemistry and biology.";
 RL Adv. Protein Chem. 26:279-402(1972).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
 RA Isaacs N.W., Agarwal R.C.;
 RT "Experience with fast Fourier least squares in the refinement of the
 RT crystal structure of rhombohedral 2-zinc insulin at 1.5-A
 RT resolution.";
 RL Acta Crystallogr. A 34:782-791(1978).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
 RX MEDLINE=89099316; PubMed=2905485;
 RA Baker E.N., Blundell T.L., Cutfield J.F., Cutfield S.M., Dodson E.J.,
 RA Dodson G.G., Crowfoot Hodgkin D.M., Hubbard R.E., Isaacs N.W.,
 RA Reynolds C.D., Sakabe K., Sakabe N., Vijayan N.M.;
 RT "The structure of 22n pig insulin crystals at 1.5-A resolution.";
 RL Philos. Trans. R. Soc. Lond., B, Biol. Sci. 319:369-456(1988).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).

RX MEDLINE=92126260; PubMed=172613;
 RA Baischmidt P., Hansen P.B., Dodson E., Dodson G., Korber P.;
 RT "Structure of porcine insulin co-crystallized with chloride Z."
 RL Acta Crystallogr. B 47:935-946(1991).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY.
 RX MEDLINE=9422450; PubMed=2025410;
 RA Badger J., Harris M.R., Reynolds C.D., Evans A., Dodson E.;
 RT "Structure of the pig insulin dimer in the cubic crystal."
 RL Acta Crystallogr. B 47:127-136(1991).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (1.65 ANGSTROMS).
 RA Diao J.-S., Wan Z.-L., Chang W.-R., Liang D.-C.;
 RT "Structure of monomeric porcine DesB1-B2 despancpeptide (E26-E30)
 RL Acta Crystallogr. D 53:507-512(1997).
 CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
 CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
 CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
 CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
 CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
 CC DISULFIDE BONDS.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
 CC -!- DATABASE: NAME=Protein Spotlight;
 CC NOTE=Issue 9 of April 2003;
 CC WWW="http://www.expasy.org/spotlight/articles/spilt_09.htm".
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is not re-used
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: AF044555; AAC77923.1; ALI_NIT.
 DR EMBL: AY044828; AAL69565.1; -.
 DR PDB: 3INS; 09-JAN-99.
 DR PDB: 4INS; 31-JUL-94.
 DR PDB: 6INS; 31-JAN-94.
 DR PDB: 7INS; 31-JAN-94.
 DR PDB: 9INS; 15-OCT-91.
 DR PDB: 11ZA; 15-OCT-91.
 DR PDB: 11ZB; 15-OCT-91.
 DR PDB: 2TC1; 29-JAN-96.
 DR PDB: IME3; 29-JAN-96.
 DR PDB: 3KTH; 29-JAN-96.
 DR PDB: 1DE1; 16-JUN-97.
 DR PDB: 1SDB; 01-APR-98.
 DR PDB: 1WAF; 28-FEB-97.
 DR PDB: 12FI; 16-FEB-99.
 DR PDB: 12NI; 28-JAN-98.
 DR PDB: 12NJ; 26-JAN-98.
 DR InterPro: IPR004825; Ins/IGF/relax.
 DR Pfam: PF00049; Insulin; 1.
 DR SMART: SM00078; IIGF; 1.
 DR PROSITE: PS00262; INSULIN; 1.
 KW Insulin family; Hormone; Glucose metabolism; Signal; 3D-structure.
 FT SIGNAL 1 24
 FT CHAIN 25 54 INSULIN B CHAIN.
 FT PROPEP 57 95 C PEPTIDE.
 FT CHAIN 88 108 INSULIN A CHAIN.
 FT DISULFID 31 94 INTERCHAIN.
 FT DISULFID 43 107 INTERCHAIN.
 FT DISULFID 93 98
 FT HELIX 26 45
 FT STRAND 48 94
 FT HELIX 99 94
 FT HELIX 103 106
 FT STRAND 107 107
 FT SEQUENCE 108 AA: 11071 MF: CB4491H4298588DE CRC64;

Query Match 82.7%; Score 383; DB 1; Length 108;
 Best Local Similarity 86.0%; Pred. No. 3.3e-34;
 Matches 74; Conservative 1; Mismatches 9; Indels 2; Gaps 1;
 QY 1 FVNHLCGSHLYVLCVGERGFYTPKTRREAESLQVQVHGGGPGAGSLQPIALHG 60
 LD 25 FVNHLCGSHLYVLCVGERGFYTPKTRREAENPAGAVELGG--GLGGIQLALEG 82
 QY 61 SLQKRGIVECCCTSLCYOLENYCN 86
 LD 83 PPQKRGIVECCCTSLCYOLENYCN 108
 RESULT 15
 INS_PSAOB STANDARD; PRT; 110 AA.
 ID INS_PSAOB
 AC Q625R7;
 DT 31-NOV-1997 (Rel. 25, Created)
 DT 31-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1998 (Rel. 35, Last annotation update)
 DE Insulin precursor.
 GN INS
 OS Psammomys obesus.
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;
 OC Psammomys.
 OC NCBI_TaxID=48139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RX MEDLINE=97305250; PubMed=9166665;
 RA Kaiser N., Baillyes E.M., Schneider B.S., Cerasi E., Steiner D.F.,
 RA Burton J.C., Gross D.C.;
 RT "Characterization of the unusual insulin of Psammomys obesus, a
 RT rodent with nutrition-induced NIDDM-like syndrome."
 RL Diabetes 46:953-957(1997).
 CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
 CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
 CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
 CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
 CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
 CC DISULFIDE BONDS.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is not re-used
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: X98241; CAA66897.1; -.
 DR HSP; P01328; IAIU.
 DR InterPro: IPR004825; Ins/IGF/re-ax.
 DR Pfam: PF00049; Insulin; 1.
 DR PROSITE: PS00262; INSULIN; 1.
 DR SMART: SM00078; IIGF; 1.
 DR PROSITE: PS00262; INSULIN; 1.
 KW Insulin family; Hormone; Glucose metabolism; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 54 INSULIN B CHAIN.
 FT PROPEP 57 87 C PEPTIDE.
 FT CHAIN 90 110 INSULIN A CHAIN.
 FT DISULFID 31 96 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 43 109 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 95 100 BY SIMILARITY.
 FT SEQUENCE 110 AA: 12324 MW: A006738E20579CB0 CRC64;
 Query Match 81.4%; Score 377; DB 1; Length 110;
 Best Local Similarity 81.4%; Pred. No. 1.5e-33;


```

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 20, Last sequence update)
DE Insulin.
OS Spermophilus tridecemlineatus (Thirteen-lined ground squirrel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Scuriinae;
OC Spermophilus.
OX NCBI_TaxID=43179;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas.
RA Tredrea M.M., Buck M.J., Guhaniyogi J., Squire T.L., Andrews M.T.;
RT "Regulation of PDK4 expression in a hibernating mammal."
RL Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
CC 1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC 1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: A018664; AAK7258.1;
DR HSSP: P01368; 11NF.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin; 1.
DR PRINTS: PR00277; INSULIN.
DR SMART: SM00078; IIGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
SQ SEQUENCE 110 AA: 12069 MW: 451769D6622BEE5 CRC64:
      Query Match: 89.2%; Score 419; DB 11; Length 110;
      Best Local Similarity 89.0%; Pred. No. 3; 4e-41;
      Matches 77; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVYVCGGSGFYTPKTRREAEDLVQGVVGLGGPGAGSLQPLALEGSLQKRG 60
DE 25 FVNHLCGSHLVYVCGGSGFYTPKTRREAEDLVQGVVGLGGPGAGSLQPLALEGSLQKRG 45
QY 61 SLQKRGIVVCGGTSICSLVQLENYCN 86
DE 85 PLQKRGIVVCGGTSICSLVQLENYCN 110

RESULT 3
Q8WNW6 PRELIMINARY; PRT; 119 AA.
AC Q8WNW6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 21, Last annotation update)
DE Preproinsulin.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas.
RA Okamoto S., Marimatsu M.;
RT "cat. insulin."
RL Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
CC 1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC 1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: A043535; BAB8410.1;
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin; 1.
DR PRINTS: PR00277; INSULIN.
DR SMART: SM00078; IIGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
SQ SEQUENCE 110 AA: 12069 MW: 95PB6E17075E7A4 CRC64:
      Query Match: 83.9%; Score 388; DB 6; Length 110;
      Best Local Similarity 83.7%; Pred. No. 3; 1e-38;
      Matches 72; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

QY 1 FVNHLCGSHLVYVCGGSGFYTPKTRREAEDLVQGVVGLGGPGAGSLQPLALEGSLQKRG 60
DB 25 FVNHLCGSHLVYVCGGSGFYTPKTRREAEDLVQGVVGLGGPGAGSLQPLALEGSLQKRG 64

```

```

QY 61 SLQKRGIVVCGGTSICSLVQLENYCN 86
DE 85 PLQKRGIVVCGGTSICSLVQLENYCN 110

RESULT 4
Q8H780 PRELIMINARY; PRT; 65 AA.
AC Q8H780;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Insulin (Fragment).
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RA O'Sullivan C., Tichy H., Klein J.;
RT "Molecular evolution in higher primates: gene specific and organism
RT specific characteristics."
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL: AY092023; AAM76640.1;
DR NON_TER 1;
DR NON_TER 65;
SQ SEQUENCE 65 AA: 6920 MW: B772017FD8BCABEA CRC64:
      Query Match: 73.9%; Score 342; DB 6; Length 65;
      Best Local Similarity 100.0%; Pred. No. 4; 7e-33;
      Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 CGSHLVYVCGGSGFYTPKTRREAEDLVQGVVGLGGPGAGSLQPLALEGSLQKRG 66
DE 1 CGSHLVYVCGGSGFYTPKTRREAEDLVQGVVGLGGPGAGSLQPLALEGSLQKRG 60

QY 67 IVEQC 71
DE 61 IVEQC 65

RESULT 5
Q8H780 PRELIMINARY; PRT; 65 AA.
AC Q8H780;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Insulin (Fragment).
OS Longe pygmaeus (orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pongo.
OX NCBI_TaxID=9680;
RN [1]
RP SEQUENCE FROM N.A.
RA O'Sullivan C., Tichy H., Klein J.;
RT "Molecular evolution in higher primates: gene specific and organism
RT specific characteristics."
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL: AY092024; AAM76641.1;
DR NON_TER 1;
DR NON_TER 65;
SQ SEQUENCE 65 AA: 6920 MW: B772017FD8BCABEA CRC64:
      Query Match: 73.9%; Score 342; DB 6; Length 65;
      Best Local Similarity 100.0%; Pred. No. 4; 7e-33;
      Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 CGSHLVYVCGGSGFYTPKTRREAEDLVQGVVGLGGPGAGSLQPLALEGSLQKRG 66
DE 1 CGSHLVYVCGGSGFYTPKTRREAEDLVQGVVGLGGPGAGSLQPLALEGSLQKRG 60

```



```

RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AF373021; AAK5558.1; -.
DR HSSP: P01308; ILNP.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin; 1
DR PRINTS: PR00277; INSULIN.
DR SMART: SM00078; IIGF; 1.
DR PROSITE: PS00262; INSULIN.
SQ SEQUENCE 108 AA; 11861 MW; D713024E22E2F5D59 CRC64;

Query Match 45.9%; Score 210.5; DP 13; Length 108;
Best Local Similarity 50.6%; Pred. No. 2; Gap 3;
Matches 43; Conservative 11; Mismatches 28; Indels 2; Gaps 2;

QY 4 QHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQLGGGAGSLQFLALEGS- 61
DB 27 QHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQLGGGAGSLQFLALEGS- 61
QY 62 LQKRGIVEQCCTSTCSLYQLENYCN 86
DB 73 ADFAFADHAEVINKRGVQGGCHKPCSFELQNYCN 108

RESULT 13
Q98TA9 ID Q98TA9 PRELIMINARY: PRT; 57 AA.
AC Q98TA9;
CT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DI 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Preproinsulin (Fragment).
OS Gnathemus petensis.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Mormyridae; Gnathemidae.
CX NCBI_TaxID=42645;
RN [1]
RP MEDLINE=2103577; PubMed=1306171;
RA Al-Mahrouki A.A., Irwin D.M., Graham L.C., Youson J.H.;
RT "Molecular cloning of preproinsulin cDNAs from several
RT osteoglossomorphs and a cyprinid."
RL Mol. Cell. Endocrinol. 174:51-56(2001).
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AF199587; AAK2871.1; -.
DR HSSP: P01308; ILNP.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin; 1
DR SMART: SM00078; IIGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
FT NON_TER 87
FT NON_TER 87
SQ SEQUENCE 87 AA; 9874 MW; FF448ED35E2453B5 CRC64;

Query Match 45.9%; Score 210.5; DP 13; Length 108;
Best Local Similarity 50.6%; Pred. No. 2; Gap 3;
Matches 43; Conservative 11; Mismatches 28; Indels 2; Gaps 2;

QY 4 QHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQLGGGAGSLQFLALEGS- 61
DB 27 QHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQLGGGAGSLQFLALEGS- 61
QY 62 LQKRGIVEQCCTSTCSLYQLENYCN 86
DB 73 ADFAFADHAEVINKRGVQGGCHKPCSFELQNYCN 108

RESULT 14
Q98TB1 ID Q98TB1 PRELIMINARY: PRT; 108 AA.
AC Q98TB1;
CT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DI 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Preproinsulin (Fragment).
OS Gnathemus petensis.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Mormyridae; Gnathemidae.
CX NCBI_TaxID=42645;
RN [1]
RP MEDLINE=2103577; PubMed=1306171;
RA Al-Mahrouki A.A., Irwin D.M., Graham L.C., Youson J.H.;
RT "Molecular cloning of preproinsulin cDNAs from several
RT osteoglossomorphs and a cyprinid."
RL Mol. Cell. Endocrinol. 174:51-56(2001).
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AF199587; AAK2871.1; -.
DR HSSP: P01308; ILNP.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin; 1
DR SMART: SM00078; IIGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
FT NON_TER 87
FT NON_TER 87
SQ SEQUENCE 87 AA; 9874 MW; FF448ED35E2453B5 CRC64;

Query Match 45.9%; Score 210.5; DP 13; Length 108;
Best Local Similarity 50.6%; Pred. No. 2; Gap 3;
Matches 43; Conservative 11; Mismatches 28; Indels 2; Gaps 2;

QY 4 QHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQLGGGAGSLQFLALEGS- 61
DB 27 QHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQLGGGAGSLQFLALEGS- 61
QY 62 LQKRGIVEQCCTSTCSLYQLENYCN 86
DB 73 ADFAFADHAEVINKRGVQGGCHKPCSFELQNYCN 108

RESULT 15
Q98TB2 ID Q98TB2 PRELIMINARY: PRT; 91 AA.
AC Q98TB2;
CT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DI 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Preproinsulin (Fragment).
OS Ambloplites rupestris (Rock bass).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Percoidae;
OC Centrarchidae; Ambloplites.
CX NCBI_TaxID=109273;
RN [1]
RP MEDLINE=109273; PubMed=1306171;
RA Al-Mahrouki A.A., Irwin D.M., Youson J.H.;
RT "Molecular cloning of preproinsulin cDNA from the rock bass."
RT Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AF199584; AAK28708.1; -.
DR HSSP: P01308; ILNP.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin; 1.
DR SMART: SM00078; IIGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
FT NON_TER 91
FT NON_TER 91
SQ SEQUENCE 91 AA; 10100 MW; E86C8B256DC69D39 CRC64;

Query Match 44.0%; Score 203.5; DP 13; Length 91;

```


Claim 12: Page 29 35: 46pp: English.

This sequence represents a human insulin precursor comprising insulin A and B chains. This insulin precursor is a component of the chimeric proteins HGH-mini-proinsulin (AA14286C) and the chimeric protein given in AA4286C. These chimeric proteins additionally contain an N-terminal fragment of human growth hormone (hGH) and a cleavable peptide linker (AA4286S). The hGH portion of the chimeric protein acts as an intramolecular chaperone (IMC) for the insulin precursor, enabling it to fold correctly. The cleavable peptide linker has a C-terminal Arg residue which enables the hGH portion of the chimeric protein to be removed after folding has taken place. Production of recombinant human insulin via an hGH-pro-insulin chimeric protein can provide human insulin with correctly linked cysteine bridges with fewer necessary procedural steps, and hence resulting in a higher yield of human insulin. The IMC sequences not only protect insulin sequences from intracellular degradation by a microorganism host, but also promote the folding of the fused insulin precursor, facilitate the solubility of the fusion protein and decrease the intermolecular interactions among the fusion proteins, thus allowing folding of the fused insulin precursor at commercially useful high concentrations. The procedural steps of cyanogen bromide cleavage, oxidative sulphydrololysis and related purification steps can thus be eliminated, along with the use of high concentrations of mercaptan or the use of hydrophobic absorbent resins.

Sequence 52 AA:

```

XX      Pro-insulin is produced by treating recombinant precursor protein
CC      with a mercaptan to provide 2-10 SH residues per Cys residue, in
CC      presence of a chaotropic agent and in aq. medium of pH 10-11,
CC      treating the prod. with 3-50 q hydrophobic adsorbent resin per 1 aq.
CC      medium of pH 4-7, isolating the adsorbed resin and pro-insulin and
CC      desorbing the pro-insulin. This method produces pro-insulin with
CC      correctly bonded Cys bridges. Compared with known methods it
CC      involves fewer stages (esp. no sulphitolysis or cyanogen bromide
CC      cleavage) and overall losses during purification are reduced, i.e.
CC      the process is quicker and gives better yields.
CC      Sequences of insulin chain A, B and C are given in AA68895-97.
CC      Sequences of pro-insulin 1-4 are given in AA68898-901.
CC      (Updated on 25-MAR-2003 to correct PN field.)
XX
XX      Sequence      56 AA;
XX
XX      Query Match          100.0%; Score 294; DB 15; Length 56;
XX      Best Local Similarity 100.0%; Pred No. 1.7e+26;
XX      Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps
XX
Oy      1 FYNQHLGSHVFAEALYLVCGERGEFYTPKTRGVVEQCSTICSJLYOLENYCN 52
Da      |||||||.....|
Da      5 FYNQHLGSHVFAEALYLVCGERGEFYTPKTRGVVEQCSTICSJLYOLENYCN 56
XX
RESULT 3
ID      AAR78665 standard; Protein; 56 AA.
AC      AAR78665;
XX
PE      03-APR-1996 (first entry)
XX
XX      Proinsulin sequence 3.
XX
XX      Proinsulin; post-translational modification: recombinant production;
KW      Protein folding; conformation.
XX
XX      Synthetic.
XX
FH      Key              Location/Qualifiers
FT      Region          1..4
FT      FT              /label= R2
FT      FT              /note= "a peptide of 4 amino acids"
FT      FT              5..14
FT      FT              /label= R1-(B2-B29)-Y
FT      FT              /note= "human insulin B-chain"
FT      FT              35
FT      FT              /label= X
FT      FT              35..56
FT      FT              /label= Gly-(A2-A20)-R3
FT      FT              /note= "human insulin A-chain"
XX
XX      EF664292-A2.
XX
XX      Z3-AGO-1995.
XX
XX      09-FEB-1995;    9589-Q1Q1746.
XX
XX      18-FEB-1994;    94DE-4405179.
XX
XX      (FARR ) HORGHST AG.
XX
XX      Geri M, Ludwig J, Obermeyer R, Sabel W;
XX
XX      WP1: 1995-284754/38.
XX
XX      Isolation of insulin that is correctly post-translationally
PI      processed - by reacting pro-insulin with a mercaptan in the presence
PI      of a chaotropic agent and purificn. after absorption to hydrophobic
PI      resin

```


Matches 52: Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY 1 FVNQHLGSHLVEALYLIVCGRGPFTPKTRGIVEQCCTSIQSYOLENYCN 52
 III III III III III III III III III III III III III III III
 DB 45 FVNQHLGSHLVEALYLIVCGRGPFTPKTRGIVEQCCTSIQSYOLENYCN 96

RESULT 6
 AAR78662
 ID AAR78662 standard; protein: 96 AA.
 XX
 AC AAR78662;
 DT 03-APR-1996 (first entry)
 XX
 DE Fusion protein contg. proinsulin sequence 3.
 XX
 KW Proinsulin; post-translational modification; recombinant production;
 KW protein folding; conformation.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Region 41..44
 FT /label= R2
 FT /note= "a peptide of 4 amino acids"
 FT Peptide 45..74
 FT /label= R1-(B2-B29)-Y
 FT /note= "human insulin B-chain"
 FT Region 75
 FT /label= X
 FT Peptide 76..96
 FT /label= G1Y (A2-A20)-Y3
 FT /note= "human insulin A-chain"
 XX
 PN EP668292-A2.
 XX
 XX 23-AUG-1995.
 PD
 PF 09-FEB-1995; 95RP-0101743.
 XX
 PR 18-FEB-1994; 44DF-4405174.
 XX
 PA (FARM) : ROCHST AG.
 XX
 PI Gerl M, Ladwig G, Obermeier R, Sabel W;
 XX
 XX WPI: 1995-254754/35.
 DR
 XX
 XX Isolation of insulin that is correctly post-translationally
 PT processed - by reacting proinsulin with a mercaptan in the presence
 PI of a chaotropic agent and purification after absorption to hydrophobic
 PI resin
 XX
 XX Example 2: Page 6; 16pp; German.
 PS
 CC The present sequence is that of a fusion protein, produced in E. coli
 CC which contains an example of a proinsulin molecule corresp.
 CC to the general formula R2-B1-(B2-B29)-Y-X-Gly-(A2-A20)-R3 (11). In
 CC formula (11), X = Lys, Arg or a peptide of 2-35 amino acids contg.
 CC Lys or Arg at the N- and C-termini; Y = a natural amino acid; R1 =
 CC Phe or a bond; R2 = H, Arg, Lys, a peptide of 2-45 amino acids
 CC contg. Arg or Lys at the N- and C-termini; R3 = a natural amino
 CC acid; (A2-A20) and (B2-B29) are the insulin A- and B-chain sequences
 CC from human or other insulin. The proinsulin molecule, released by
 CC cyanogen bromide, is reacted with mercaptan at a ratio of 2-10 SH
 CC residues of mercaptan per Cys residue of proinsulin. The reaction
 CC takes place in the presence of a chaotropic auxiliary agent at
 CC pH 10-11 and results in proinsulin with correctly linked cysteine
 CC bridges. Reaction with trypsin and opt. carboxypeptidase B yields
 CC correctly folded insulin. The insulin is isolated by absorption on
 CC a hydrophobic resin.

Seq Sequence 96 AA;
 Query Match 100.0%; Score 294; DB 16; Length 96;
 Best Local Similarity 100.0%; Pred. No. 2.9e-26;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQHLGSHLVEALYLIVCGRGPFTPKTRGIVEQCCTSIQSYOLENYCN 52
 III III III III III III III III III III III III III III III
 DB 45 FVNQHLGSHLVEALYLIVCGRGPFTPKTRGIVEQCCTSIQSYOLENYCN 96

RESULT 7
 AAY42860
 ID AAY42860 standard; protein: 107 AA.
 XX
 AC AAY42860;
 XX
 DT 19-JAN-2000 (first entry)
 DT
 DE hGH-mini-proinsulin chimeric protein.
 XX
 XX
 KW Insulin; precursor; growth hormone; chaperone; intramolecular;
 KW folding; conformation; chimeric protein; cleavable; recombinant;
 KW production; yield.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO9950302-A1.
 XX
 PD 07-OCT-1999.
 XX
 PF 31-MAR-1998; 98WO-CN00352.
 XX
 PR 31-MAR-1998; 95WO-CN00652.
 XX
 PA (TONG-) TONGHUA GANTECH BIOTECHNOLOGY LTD.
 XX
 XX Gan Z;
 XX
 XX WPI: 1999-5:0835/52.
 XX
 XX New chimeric proteins containing human growth hormone fragment, used
 FT particularly for the production of human insulin
 FT
 XX
 XX Claim 13: Page 30; 46pp; English.
 XX
 CC This sequence represents a chimeric protein, hGH-mini-proinsulin.
 CC This chimeric protein contains an N-terminal fragment of human growth
 CC hormone (hGH) of the sequence given in AAY42855, a cleavable peptide
 CC linker (AAY42857), and a human insulin precursor comprising insulin
 CC A and B chains (AAY42859). The hGH portion of the chimeric protein acts
 CC as an intramolecular chaperone (IMC) for the insulin precursor,
 CC enabling it to fold correctly. The cleavable peptide linker has a
 CC C-terminal Arg residue which enables the hGH portion of the
 CC chimeric protein to be removed after folding has taken place. Production
 CC of recombinant human insulin via an hGH-proinsulin chimeric protein can
 CC provide human insulin with correctly linked cysteine bridges with
 CC fewer necessary procedural steps, and hence resulting in a higher yield
 CC of human insulin. The IMC sequences not only protect insulin sequences
 CC from intracellular degradation by a microorganism host, but also promote
 CC the folding of the fused insulin precursor, facilitate the solubility of
 CC the fusion protein and decrease the intermolecular interactions among
 CC the fusion proteins, thus allowing folding of the fused insulin precursor
 CC at commercially useful high concentrations. The procedural steps of
 CC cyanogen bromide cleavage, oxidative sulphylation and related
 CC purification steps can thus be eliminated, along with the use of high
 CC concentrations of mercaptan or the use of hydrophobic absorbent resins.

Seq Sequence 107 AA;
 Query Match 100.0%; Score 294; DB 20; Length 107;
 Best Local Similarity 100.0%; Pred. No. 3.2e-26;

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQLCGSHLVEALYLVCGERGFFYTPKTRGLIVEGQCTISCLSYLQENYCN 12
 DB 56 FVNQLCGSHLVEALYLVCGERGFFYTPKTRGLIVEGQCTISCLSYLQENYCN 137

RESULT 8
 AAR8897
 ID AAR9897 standard; Protein: 116 AA.
 AC AAR8897;
 DT 03-FEB-1997 (first entry)
 XX SQB-proinsulin hybrid polypeptide.
 XX Insulin; proinsulin hybrid polypeptide; protein foldings;
 KW enzymatic cleavage; cyanogen bromide; sulphydrolysts.
 KW Homo sapiens.
 OS
 XX MO9620724-A1.
 PN 11-JUL-1996.
 PD 29-DEC-1994; 94WO-05132#3.
 PF 29-DEC-1994; 94WO-05132#3.
 PR 10-JAN-1995; 95ZA-00001#2.
 XX (BIO-) BIO-TECHNOLOGY GENERAL CORP.
 PA Gorecki K, Hartman JR, Mendelovitz S;
 PI WPI: 1996-333766/33.
 DK N-PSDB; AAT34670.
 XX Recombinant insulin prodr. by correctly folding pro-insulin hybrid
 PT polypeptide then enzymatic cleavage of folded product, does not
 PT require sulphite protection of SH nor use of cyanogen bromide
 XX Example 1B; Figure 7; 69pp; English.
 PS
 XX A new method for the production of recombinant human insulin
 CC comprises folding a hybrid polypeptide comprising proinsulin under
 CC conditions that permit correct disulphide bond formation and
 CC subjecting that folded protein to enzymatic cleavage. The insulin
 CC produced can then be purified. This sequence is a SQB-insulin B
 CC chain-Arg-insulin A chain hybrid polypeptide and is encoded by
 CC the plasmid construct pDBAST-LAT. Transformation of the proper
 CC E.coli host cells with pDBAST-LAT results in the efficient
 CC expression of the proinsulin hybrid polypeptide. Useful for human
 CC insulin production. The method produces recombinant human insulin
 CC identical to the natural hormone. Hazardous and embryonic
 CC procedures involving cyanogen bromide and sulphydrolysts to protect
 CC SH groups are avoided since the entire hybrid polypeptide folds
 CC efficiently to the native structure even with the leader attached
 CC and Cys unprotected.
 XX
 SQ Sequence 116 AA;
 Query Match 100.0%; Score 294; DB 17; Length 116;
 Best Local Similarity 100.0%; Pred. No. 3.5e-26;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQLCGSHLVEALYLVCGERGFFYTPKTRGLIVEGQCTISCLSYLQENYCN 12
 DB 56 FVNQLCGSHLVEALYLVCGERGFFYTPKTRGLIVEGQCTISCLSYLQENYCN 137

RESULT 9
 AAR71692
 ID AAR71692 standard; Protein: 137 AA.
 AC AAR71692;
 DT 25-MAR-2003 (updated)
 DT 20-NOV-1995 (first entry)
 XX Mating factor alpha 1-insulin precursor ArgB31.
 XX Human insulin precursor ArgB31; diabetes; zinc ion complex;
 KW mating factor alpha 1.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 XX Protein 1..85
 XX /label= mating factor: alpha-1
 XX Peptide 96..116
 XX /label= B-chain
 XX Peptide 117..137
 XX /label= A-chain
 XX MO9507331-A1.
 PN 23-MAR-1995.
 PD 16-SEP-1994; 94WO-090347.
 PF 17-SEP-1994; 93DK-0001344.
 PR 02-FEB-1994; 94US-0190829.
 XX (NVO) NOVO-NORDISK AS.
 PA Anderson AS, Halstrom JB, Havelund S, Jonassen I;
 PI Markussen J;
 XX WPI: 1995-131314/17.
 DK N-PSDB; AAQ86425.
 XX Acylated insulin deriv. which may be present as a zinc ion
 PT complex - is used to treat diabetes and is rapid acting.
 XX Example 5; Page 78; 100pp; English.
 PS
 XX AAO86425 encodes AAR71692 mating factor alpha 1-insulin precursor
 CC ArgB31. ArgB31 comprises the B and A chains of a claimed human
 CC insulin derivative. In the final claimed compsn. they are covalently
 CC connected via disulphide bonds between Cys residues A7/B7 and
 CC A20/B19. The derivative, which may be present as a zinc ion
 CC complex, can be used as a fast action treatment for diabetes.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 137 AA;
 Query Match 100.0%; Score 294; DB 16; Length 137;
 Best Local Similarity 100.0%; Pred. No. 4.1e-26;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVNQLCGSHLVEALYLVCGERGFFYTPKTRGLIVEGQCTISCLSYLQENYCN 52
 DB 86 FVNQLCGSHLVEALYLVCGERGFFYTPKTRGLIVEGQCTISCLSYLQENYCN 137

RESULT 10
 AAR71694
 ID AAR71694 standard; Protein: 145 AA.
 AC AAR71694;
 XX 25-MAR-2003 (updated)
 DT 20-NOV-1995 (first entry)
 XX Mating factor alpha 1-insulin precursor ArgB1, ArgB31 N-terminal.


```

XX Human insulin precursor ArgB1, ArgB31; diabetes; zinc ion complex
KW mating factor alpha 1; N-terminal BEAEAEAR.
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT Protein: 1..85 /label= mating factor alpha 1
FT Peptide 86..93 /label= N-terminal peptide
FT Peptide 94..124 /label= B-chain
FT Peptide 125..145 /label= A-chain
XX WO9507931-A1.
XX 23-MAR-1995.
XX 16-SEP-1994; 94WO-DK00347.
XX 17-SEP-1993; 93DK-0001044.
XX 02-FEB-1994; 94US-0190829.
XX (NOVO ) NOVO-NORDISK AS.
XX Andersen AS, Halstrom JB, Havelund S, Jonassen J;
XX Markussen J;
XX WPI: 1995-131314/17.
XX N-PSDB; AAQ86423.
XX Acylated insulin deriv. which may be present as a zinc ion
XX complex - is used to treat diabetes and is rapid acting.
XX Example 5: Pages 82-83; 100pp; English.
XX AAQ86429 encodes AAR71694 mating factor alpha 1-insulin precursor
XX ArgB1, ArgB31 N-terminal BEAEAEAR. The insulin precursor comprises
XX the B and A chains of a claimed human insulin derivative preceded
XX by the N-terminal amino acids BEAEAEAR. In the final claimed compsn.
XX they are covalently connected via disulphide bonds between Cys
XX residues A7/B7 and A20/B19. The derivative, which may be present
XX as a zinc ion complex, can be used as a fast action treatment for
XX diabetes.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX Sequence 145 AA:
XX
XX Query Match 100.0%; Score 294; DB 16; Length 145;
XX Best Local Similarity 100.0%; Pred. No. 4.4e-26;
XX Matches 52; Conservative 0; Mismatches 0; Gaps 0;
XX
OY 1 FVNCHLGGSHLVEALYVGCGERRFFYPKTRGIVGQCCTSCSYOLENYCN 52
DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
94 FVNCHLGGSHLVEALYVGCGERRFFYPKTRGIVGQCCTSCSYOLENYCN 145
RESULT 11
AAR71695
ID AAR71695 standard; Protein: 146 AA.
XX
XX AAR71695:
XX
XX 25-MAR-2003 (updated)
XX 20-NOV-1995 (first entry)
XX
XX Mating factor alpha 1-insulin precursor ArgB1, ArgB31 N-terminal.
XX Human insulin precursor ArgB1, ArgB31; diabetes; zinc ion complex;
KW mating factor alpha 1; N-terminal BEAEAEAR.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Protein: 1..85 /label= mating factor alpha 1
XX Peptide 86..93 /label= N-terminal peptide
XX Peptide 94..124 /label= B-chain
XX Peptide 125..145 /label= A-chain
XX WO9507931-A1.
XX 23-MAR-1995.
XX 16-SEP-1994; 94WO-DK00347.
XX 17-SEP-1993; 93DK-0001044.
XX 02-FEB-1994; 94US-0190829.
XX (NOVO ) NOVO-NORDISK AS.
XX Andersen AS, Halstrom JB, Havelund S, Jonassen J;
XX Markussen J;
XX WPI: 1995-131314/17.
XX N-PSDB; AAQ86423.
XX Acylated insulin deriv. which may be present as a zinc ion
XX complex - is used to treat diabetes and is rapid acting.
XX Example 5: Pages 82-83; 100pp; English.
XX AAQ86429 encodes AAR71694 mating factor alpha 1-insulin precursor
XX ArgB1, ArgB31 N-terminal BEAEAEAR. The insulin precursor comprises
XX the B and A chains of a claimed human insulin derivative preceded
XX by the N-terminal amino acids BEAEAEAR. In the final claimed compsn.
XX they are covalently connected via disulphide bonds between Cys
XX residues A7/B7 and A20/B19. The derivative, which may be present
XX as a zinc ion complex, can be used as a fast action treatment for
XX diabetes.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX Sequence 146 AA:
XX
XX Query Match 100.0%; Score 294; DB 16; Length 146;
XX Best Local Similarity 100.0%; Pred. No. 4.4e-26;
XX Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 FVNCHLGGSHLVEALYVGCGERRFFYPKTRGIVGQCCTSCSYOLENYCN 52
DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111
95 FVNCHLGGSHLVEALYVGCGERRFFYPKTRGIVGQCCTSCSYOLENYCN 146
RESULT 12
AAY42861
ID AAY42861 standard; protein: 150 AA.
XX
XX AAY42861:
XX
XX 19-JAN-2000 (first entry)
XX
XX Chimeric protein, SEQ ID 7.
XX
XX Insulin; precursor; growth hormone; chaperone; intramolecular;
XX folding; conformation; chimeric protein; cleavable; recombinant;
XX production; yield.
XX
XX Synthetic.
XX Homo sapiens.
XX
XX WO950302-A1.

```

XX 07-OCT-1999.
 XX
 XX 31-MAR-1998: 98WO-CR00052.
 XX
 XX 31-MAR-1998: 98WO-CR00052.
 XX
 XX (TONG:) TONGHUA GANTHEM BIOTECHNOLOGY LTD.
 XX
 XX Gan Z:
 XX
 XX WPI: 1999 610839/52.
 XX
 XX New chimeric proteins containing human growth hormone fragment, used particularly for the production of human growth hormone.
 XX
 XX Claim 14: Page 30-31: 46pp: English.
 XX
 XX This sequence represents a chimeric protein, which contains an N-terminal fragment of human growth hormone (hGH) of the sequence given in AA42856, a cleavable peptide linker (AA42857), and a human insulin precursor comprising insulin A and B chains (AA42859). The high portion of the chimeric protein acts as an intramolecular chaperone (IMC) for the insulin precursor, enabling it to fold correctly. The cleavable peptide linker has a C-terminal Arg residue which enables the full portion of the chimeric protein to be removed after folding has taken place.
 XX
 XX Production of recombinant human insulin via an anti-proinsulin chimeric protein can provide human insulin with correctly linked cysteine bridges with fewer necessary procedural steps, and hence resulting in a higher yield of human insulin. The IMC sequences not only protect insulin sequences from intracellular degradation by a microorganism host, but also promote the folding of the fused insulin precursor, facilitate the solubility of the fusion protein and decrease the intermolecular interactions among the fusion proteins, thus allowing folding of the fused insulin precursor at commercially useful high concentrations. The procedural steps of cyanogen bromide cleavage, oxidative sulphydralysis and related purification steps can thus be eliminated, along with the use of high concentrations of mercaptan or the use of hydrophobic absorbent resins.
 XX
 XX Sequence 150 AA:
 SQ
 Query Match 100.0%; Score 294; DB 20; Length 150;
 Best Local Similarity 100.0%; Pred. No. 4.5e-26;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FVNHLCGSHLYEALYLVCGERGFFYTPKTRGIVEQCCTSCISLYOLENYCN 52
 DB 99 FVNHLCGSHLYEALYLVCGERGFFYTPKTRGIVEQCCTSCISLYOLENYCN 150
 RESULT 13
 AAR04582
 ID AAR04582 standard; protein: 57 AA.
 XX
 XX AAR04582;
 XX
 XX 25-MAR-2003 (updated)
 XX 14-SEP-1990 (first entry)
 XX
 XX Proinsulin analogue with a Lys residue linking the A and B chains.
 XX
 XX insulin fusion protein; pre-insulin analogue; tandemistore;
 KW Lys-Lys bridge: ds
 XX
 XX synthetic.
 OS
 XX
 XX Key Location/Qualifiers
 PH
 XX misc-difference 36
 FT /label-Lys residue linking insulin B chain to A chain
 FT Peptide 1..35
 FT /label- Insulin B chain
 FT Peptide 37..57

FT
 XX EP367183-A.
 XX
 XX 09-MAY-1990.
 XX
 XX 22-OCT-1989: 89EP-0120056.
 XX
 XX 03-NOV-1988: 88CE-1837273.
 XX 19-AUG-1989: 89DE-3927449.
 XX
 XX (FARH) ROECHST AG.
 XX
 XX Kol-er SP, Riess GJ, Uhlmann E, Wallmeier H;
 XX WPI: 1990-141149/13.
 XX N-PSDB: AAQ04335.
 XX
 XX New insulin fusion proteins comprise pro-insulin analogue linked to tandemistate
 XX
 XX Disclosure; Page 2: 2pp: Gorman.
 XX
 XX This sequence is joined to the C-terminus of an N-terminal fragment comprising opt. modified tandemistate. This fusion protein may be converted into human insulin using known methods. The synthetic gene was prepared by the phosphoramidite method.
 XX
 XX See also AAQ04335.
 XX (Updated on 25-MAR-2003 to correct PR field.)
 XX (Updated on 25-MAR-2003 to correct PI field.)
 XX
 XX Sequence 57 AA:
 SQ
 Query Match 99.08; Score 291; DB 11; Length 57;
 Best Local Similarity 98.1%; Pred. No. 3.8e-26;
 Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FVNHLCGSHLYEALYLVCGERGFFYTPKTRGIVEQCCTSCISLYOLENYCN 52
 DB 6 FVNHLCGSHLYEALYLVCGERGFFYTPKTRGIVEQCCTSCISLYOLENYCN 57
 RESULT 14
 AAR11899
 ID AAR11899 standard; Protein: 52 AA.
 XX
 XX AAR11899;
 XX
 XX 25-MAR-2003 (updated)
 XX 22-JUL-1991 (first entry)
 XX
 XX Example of human insulin precursor.
 XX
 XX Human insulin; diabetes; transpeptidation.
 XX
 XX BORO suppliers.
 XX
 XX EP427296-A.
 XX
 XX 15-MAY-1991.
 XX
 XX 29-MAY-1985: 90EP-0121887.
 XX
 XX 15-NOV-1990: 90SP-0121887.
 XX 30-MAY-1984: 84DK-0002665.
 XX 08-FEB-1985: 85DK-0000582.
 XX
 XX (NWO) NOVO-NORDISK AS.
 XX
 XX Markussen J, Fjell N, Asmerer G, Hansen MT, Thim L;
 XX Norris K. Volgt HO;
 XX WPI: 1991-141828/20.

/label- Insulin A chain

```

XX human insulin precursors - expressed with correctly positioned
PT disulphide bridges giving improved resistance to proteolysis
XX
XX Claim 3: page 18: 28pp: English.
XX
CC This human insulin precursor has correctly positioned disulphide
CC bridges between the A and B chains and is more resistant to
CC proteolytic digestion than prior art insulin precursors. Yeast
CC strains transformed with DNA encoding this precursor can be
CC cultured to secrete it in high yields. The precursor can be
CC converted into mature human insulin by transpeptidation.
CC See also AAR11897-98.
CC (Updated on 25-MAR-2003 to correct PF field.)
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 52 AA:
      Query Match          97.6%; Score 287; DB 12; Length 52;
      Best Local Similarity 95.2%; Pred. No. 9, 5e-26;
      Matches 50; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 FVNHLCGSHLVEALYLVGCGRGEFFYTPKTRIGIVCQCTSCSYOLENYCN 52
DB 1 FVNHLCGSHLVEALYLVGCGRGEFFYTPKTRIGIVCQCTSCSYOLENYCN 52
      : FVNHLCGSHLVEALYLVGCGRGEFFYTPKTRIGIVCQCTSCSYOLENYCN 52

RESULT 15
AAR65883
ID AAR65883 standard; protein; 53 AA.
XX
XX AAR65883;
XX
XX 25-MAR-2003 (updated)
XX 26-JUN-1995 (first entry)
XX
XX Di-Ary-(831-32)-human insulin amorphous, monospherical deriva.
XX Human insulin; recombinant production; amorphous; monospherical form;
XX diabetes mellitus.
XX
XX Homo sapiens (produced recombinantly in Escherichia coli).
XX
XX Key Location/Qualifiers
XX Protein 1..30
XX /label= insulin_B-chain
XX Protein 33..53
XX /label= insulin_A chain
XX
XX EP622376-A1.
XX
XX 02-NOV-1994.
XX
XX 27-APR-1994; 94EP-0106196.
XX
XX 27-APR-1993; 93BE-4313702.
XX
XX (FARH ) HCECHST AG.
XX
XX Deil P, Geisen X, Obermeier R, Sabel W;
XX WPI: 1994-334579/42.
XX
XX Amorphous, mono-spherical form of insulin deriva. - for treating
XX diabetes mellitus, are produced by diluting soln. in aq.
XX isopropanol, are stable when dried or in suspension
XX
XX Example 2: Page 5; 10pp; German.
XX
XX This sequence is a specific example of an insulin derivative which
XX can be obtained in amorphous, monospherical form by dissolving in
XX an n-propanol/buffer mixture (pH 4.5-6.5) having n-propanol content
XX 15% relative to water. The solution is then diluted with water to

```

```

CC reduce n-propanol content to below 15%. The resulting insulin
CC preparation is stable and can be used for the treatment of diabetes
CC mellitus.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 53 AA:
      Query Match          96.4%; Score 283.5; DB 15; Length 53;
      Best Local Similarity 98.1%; Pred. No. 2.5e-25;
      Matches 52; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
OY 1 FVNHLCGSHLVEALYLVGCGRGEFFYTPKTRIGIVCQCTSCSYOLENYCN 52
DB 1 FVNHLCGSHLVEALYLVGCGRGEFFYTPKTRIGIVCQCTSCSYOLENYCN 53
      : FVNHLCGSHLVEALYLVGCGRGEFFYTPKTRIGIVCQCTSCSYOLENYCN 53

Search completed: September 15, 2003, 12:00:59
Job time : 37.5305 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Computer Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2003 11:57:40 : Search time 11.9245 seconds
(without alignments)
164.449 Million cell updates/sec

Title: US-09-423-100-5

Perfect score: 294

Sequence: 1 FVNHGSSHLVDAALVCG.....LVKCCISLDSLYDENYK 52

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 3.5

Searched: 328717 seqs, 4210658 residues

Total number of hits satisfying chosen parameters: 42817

Minimum DB seq length: 0
Maximum DB seq length: 260000000

Post-processing: Minimum Match 1%
Maximum Match 100%
Listing first 15 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2_6/pdata1/1/aa/5A.COMB.pept.*
- 2: /cgn2_6/pdata1/1/aa/5B.COMB.pept.*
- 3: /cgn2_6/pdata1/1/aa/6A.COMB.pept.*
- 4: /cgn2_6/pdata1/1/aa/6B.COMB.pept.*
- 5: /cgn2_6/pdata1/1/aa/6C.COMB.pept.*
- 6: /cgn2_6/pdata1/1/aa/6D.COMB.pept.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	294	100.0	56	1	US-08-160-376A-7
2	294	100.0	56	1	US-08-389-487-31
3	294	100.0	63	1	US-08-160-376A-6
4	294	100.0	66	1	US-08-291-040B-5
5	294	100.0	96	1	US-08-160-376A-5
6	294	100.0	96	1	US-08-389-487-8
7	294	100.0	137	1	US-09-400-256-39
8	294	100.0	137	3	US-08-975-365-39
9	294	100.0	145	1	US-08-400-256-45
10	294	100.0	145	3	US-08-975-365-45
11	294	100.0	146	1	US-08-400-256-48
12	294	100.0	146	3	US-08-975-365-48
13	291	99.0	57	1	US-08-030-732A-44
14	293.5	96.4	53	1	US-08-273-17-4
15	293.5	96.4	53	4	US-08-981-988A-42
16	278.5	94.7	51	4	US-09-477-924-3
17	278.5	94.7	51	4	US-09-723-981-3
18	278.5	94.7	51	4	US-09-723-896-3
19	277.5	94.4	53	1	US-08-233-617-3
20	277	94.2	55	3	US-08-900-574-3
21	276.5	94.0	55	3	US-08-900-574-6
22	276.5	94.0	66	3	US-08-900-574-5
23	276.5	94.0	67	4	US-08-981-988A-1
24	276.5	94.0	67	4	US-08-981-988A-5
25	276	93.9	67	3	US-08-900-574-7
26	275.5	93.7	53	3	US-09-241-853-2
27	275.5	93.7	65	1	US-08-468-674B-71

Sequence 71, Appl
Sequence 41, Appl
Sequence 41, Appl
Sequence 43, Appl
Sequence 45, Appl
Sequence 15, Appl
Sequence 15, Appl
Sequence 37, Appl
Sequence 3, Appl
Sequence 36, Appl
Sequence 19, Appl
Sequence 33, Appl
Sequence 42, Appl
Sequence 33, Appl
Sequence 42, Appl
Sequence 2, Appl
Sequence 4, Appl
Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-160-376A-7
Sequence 7, Application US/08150376A
Patent No. 5473049
GENERAL INFORMATION:
APPLICANT: Obermeyer, Ranier
APPLICANT: Gerl, Martin
APPLICANT: Ludwig, Jurgen
APPLICANT: Sabel, Walter
TITLE OF INVENTION: Process For Obtaining Proinsulin
TITLE OF INVENTION: Possessing Correctly Linked
TITLE OF INVENTION: Cystine Bridges
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenneth A. Geront, Esq.
STREET: Rt. 202-206 No. 5473049th/P.O. Box 2500
CITY: Somerville
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 08876-1258
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
COMPUTER: IBM 386
OPERATING SYSTEM: WINDOWS 3.1
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/160,376A
FILING DATE: December 1, 1993
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GE P 4240420.7
FILING DATE: December 2, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Barbara V. Muller, Esq.
REGISTRATION NUMBER: 31,287
REFERENCE/DOCKET NUMBER: HOE 92/F 384
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 231-4079
TELEFAX: (908) 231-2255
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 56 Amino Acids
TYPE: Amino Acid (AA)
TOPOLOGY: not relevant
US-08-160-376A-7

Query Match 100.0%; Score 294; DB 1; Length 56;
Best Local Similarity 100.0%; Pred. No. 4.9e-29;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

; Sequence 45, Application US/08/975365
; Patent No. 6011007
; GENERAL INFORMATION:
; APPLICANT: Havelund, Svend
; APPLICANT: Halstrom, John
; APPLICANT: Jonassen, Ib
; APPLICANT: Andersen, Asger Sloth
; APPLICANT: Markussen, Jan
; TITLE OF INVENTION: ACYLATED INSULIN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 60110070 No. 60110070disk of No. 60110070th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 03-MAR-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3985.220-US
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-975-365-45
; Query Match 100.0%; Score 294; DB 1; Length 145;
; Best Local Similarity 100.0%; Pred. No. 1.4e-28;
; Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

; Sequence 48, Application US/03404256
; Patent No. 5753497
; GENERAL INFORMATION:
; APPLICANT: Havelund, Svend
; APPLICANT: Halstrom, John
; APPLICANT: Jonassen, Ib
; APPLICANT: Andersen, Asger Sloth
; APPLICANT: Markussen, Jan
; TITLE OF INVENTION: ACYLATED INSULIN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 57534970 No. 57534970disk of No. 57534970th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 03-MAR-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3985.220-US
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 146 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-975-365-48
; Query Match 100.0%; Score 294; DB 1; Length 146;
; Best Local Similarity 100.0%; Pred. No. 1.4e-28;
; Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 03-MAR-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3985.220-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 146 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-975-365-48
; Query Match 100.0%; Score 294; DB 1; Length 146;
; Best Local Similarity 100.0%; Pred. No. 1.4e-28;
; Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

; Sequence 48, Application US/08975365
; Patent No. 6011007
; GENERAL INFORMATION:
; APPLICANT: Havelund, Svend
; APPLICANT: Halstrom, John
; APPLICANT: Jonassen, Ib
; APPLICANT: Andersen, Asger Sloth
; APPLICANT: Markussen, Jan
; TITLE OF INVENTION: ACYLATED INSULIN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 60110070 No. 60110070disk of No. 60110070th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 03-MAR-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/400,256
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3985.220-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 146 amino acids

```



```

: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-975-365-48

Query Match      100.0%  Score 291;  DB 4;  Length 146;
Best local Similarity 100.0%;  Pred. No. 1.4e-28;
Matches 52;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTGRGIVEQCCTSIQSYOLENYCN 52
DB 55 FVNQHLGSHLVEALYLVCGERGFFYTPKTGRGIVEQCCTSIQSYOLENYCN 146

RESULT 13
US-06-030-731A-44
: Sequence 44, Application US/06030731A
: Patent No. 5426036
: GENERAL INFORMATION:
: APPLICANT: Koller, Klaus-Peter
: APPLICANT: Riess, Guenther Johannes
: APPLICANT: Uhlmann, Eugen
: APPLICANT: Wallmeier, Holger
: TITLE OF INVENTION: Processes for the Preparation of Foreign
: TITLE OF INVENTION: Proteins in Streptomyces
: NUMBER OF SEQUENCES: 48
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
: ADDRESSEE: Dunner
: STREET: 1300 I Street, N.W., Suite 700
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20005-3315
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/06/030-731A
: FILING DATE: 12-MAR-1993
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/189,840
: FILING DATE: 03-MAY-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/430,622
: FILING DATE: 01-NOV-1989
: APPLICATION NUMBER: US 07/687,610
: FILING DATE: 19-APR 1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/735,759
: FILING DATE: 29-JUL-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: DE P 37 14 866.4
: FILING DATE: 05-MAY-1987
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: DE P 38 37 273.8
: FILING DATE: 03-NOV-1988
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: DE P 39 27 449.7
: FILING DATE: 19-AUG-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: DE P 40 12 812.3
: FILING DATE: 21-APR-1990
: ATTORNEY/AGENT INFORMATION:
: NAME: Kirschner, Michael; 4
: REGISTRATION NUMBER: 34,551
: REFERENCE/DOCKET NUMBER: 02481-0593-02500
: TELEPHONE: 202-408-4000
```

```

: TELEFAX: 202-408-4400
: INFORMATION FOR SEQ ID NO: 44:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 57 amino acids
: TYPE: amino acid
: TOPOLOGY: unknown
: MOLECULE TYPE: peptide
US-08-030-731A-44

Query Match      99.0%  Score 291;  DB 1;  Length 57;
Best local Similarity 98.1%;  Pred. No. 1.2e-28;
Matches 51;  Conservative 1;  Mismatches 0;  Indels 0;  Gaps 0;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTGRGIVEQCCTSIQSYOLENYCN 52
DB 6 FVNQHLGSHLVEALYLVCGERGFFYTPKTGRGIVEQCCTSIQSYOLENYCN 57

RESULT 14
US-08-233-617-4
: Sequence 4, Application US/08233617
: Patent No. 5456666
: GENERAL INFORMATION:
: APPLICANT: Overmeier, Rainer
: APPLICANT: Sabel, Walter
: APPLICANT: Deil, Peter
: APPLICANT: Geisen, Karl
: TITLE OF INVENTION: Amorphous Monospherical Forms of Insulin
: TITLE OF INVENTION: Derivatives
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
: ADDRESSEE: Dunner
: STREET: 1300 I Street, N.W., Suite 700
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20005-3315
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/233.617
: FILING DATE: 25-APR-1994
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: P 43 13 702.4
: FILING DATE: 27-APR-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Carol P. Einandi
: REGISTRATION NUMBER: 32,220
: REFERENCE/DOCKET NUMBER: 02481.1374-00000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-408-4000
: TELEFAX: 202-408-4400
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 53 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: ORIGINAL SOURCE:
: ORGANISM: Escherichia coli
US-08-233-617-4

Query Match      96.4%  Score 283.5;  DB 1;  Length 53;
Best local Similarity 98.1%;  Pred. No. 8.9e-28;
Matches 52;  Conservative 0;  Mismatches 0;  Indels 1;  Gaps 1;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTGRGIVEQCCTSIQSYOLENYCN 52
DB 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTGRGIVEQCCTSIQSYOLENYCN 52
```

DB 2 FVNQLCGSHLVEALYLVCGERGFFYTPKTRRLVHQCTSCSLYOLENTON 53

RESULT: 15

```

1  : US-08-981-989A-42
2  : Sequence 42, Application: US/04987498A
3  : Patent No. 5,837,94
4  : GENERAL INFORMATION:
5  : APPLICANT: Vittal Mallaya Scientific Research Foundation
6  : TITLE OF INVENTION: The University of Leicester
7  : TITLE OF INVENTION: Insulin
8  : NUMBER OF SEQUENCES: 4
9  : CORRESPONDENCE ADDRESS:
10 : ADDRESSEE: VITTAL MALLAYA SCIENTIFIC RESEARCH FOUNDATION
11 : STREET: K. K. ROAD
12 : CITY: BANGALORE
13 : COUNTRY: INDIA
14 : ZIP: 560 004
15 :
16 : COMPUTER READABLE FORM:
17 : MEDIUM TYPE: Floppy disk
18 : COMPUTER: IBM PC Compatible
19 : OPERATING SYSTEM: PC-DOS/MS-DOS
20 : SOFTWARE: Patent Release #1.0, Version 1.1.0 (BPG)
21 : CURRENT APPLICATION DATA:
22 : APPLICATION NUMBER: US/08/381-989A

```

Query Match	96.44	Score 283.5	LR 4	Length: 52
-------------	-------	-------------	------	------------

Query Match
Best Local Similarity
90.44; Score 263.07; L₁ 4; Length: 1
98.14; Pred. No. 3.8C-28;

Best local similarity: 50.14%, frequency: 0.0025,
Matches 52; Conservative 0; Mismatches 3; Indels 1; Gaps 1

Qy 1 FVNGHLCGSHLVEALYIVCGERGFYTPKT-RGVFQCGTSCSLNLIENYCN "E

QY
1 FVNRHLCGSHLVEALIVCGDGFYYTKTKRGIVFGCSGSGLSYLVENYN
.....
QZ
1 FVNQHLCGSHLVEALIVCGDGEFFYYTKTKRGIVFGCSGSGLSYLVENYN

Search completed: September 15, 2003, 12:05:37
Job time : 11.9287 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw mode:

Run on: September 15, 2003, 12:03:35 : Search time 21.8065 Seconds
(without alignments)
347,945 Million cell updates/sec

Title: us-09-423-100-5

Perfect score: 294

Sequence: 1 FVNQRLCGSHVAFIVCG.....IVEQCGSHQSHQVNCIENTCN 52

Scoring table: BLOSUM62

Gapop 16.0 , Gapext 0.5

Searched: 54,936 seqs, 1459,2426 residues

Total number of hits satisfying chosen parameters: 54136

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications_AA:

- 1: /cgn2_6/ptcddata/1/pubaa/us07_PUBCOMB.pep
- 2: /cgn2_6/ptcddata/1/pubaa/PCT_NEW_PUB.pep
- 3: /cgn2_6/ptcddata/1/pubaa/US06_NEW_PUB.pep
- 4: /cgn2_6/ptcddata/1/pubaa/US06_PUBCOMB.pep
- 5: /cgn2_6/ptcddata/1/pubaa/US07_NEW_PUB.pep
- 6: /cgn2_6/ptcddata/1/pubaa/PCT_NEW_PUBCOMB.pep
- 7: /cgn2_6/ptcddata/1/pubaa/US06_NEW_PUB.pep
- 8: /cgn2_6/ptcddata/1/pubaa/US08_PUBCOMB.pep
- 9: /cgn2_6/ptcddata/1/pubaa/US09_PUBCOMB.pep
- 10: /cgn2_6/ptcddata/1/pubaa/US09_PUBCOMB.pep
- 11: /cgn2_6/ptcddata/1/pubaa/US09_PUBCOMB.pep
- 12: /cgn2_6/ptcddata/1/pubaa/US09_NEW_PUB.pep
- 13: /cgn2_6/ptcddata/1/pubaa/US10_PUBCOMB.pep
- 14: /cgn2_6/ptcddata/1/pubaa/US10_PUBCOMB.pep
- 15: /cgn2_6/ptcddata/1/pubaa/US10_PUBCOMB.pep
- 16: /cgn2_6/ptcddata/1/pubaa/US12_NEW_PUB.pep
- 17: /cgn2_6/ptcddata/1/pubaa/US06_NEW_PUB.pep
- 18: /cgn2_6/ptcddata/1/pubaa/US06_PUBCOMB.pep

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	294	100.0	52	14	US-10-054-873-5
2	294	100.0	107	14	US-10-054-873-5
3	294	100.0	150	14	US-10-054-873-5
4	278.5	94.7	51	11	US-09-658-935R-5
5	278.5	94.7	51	14	US-10-028-410-3
6	275.5	93.7	124	10	US-09-894-711-13
7	275.5	93.7	138	10	US-09-861-687-19
8	273	92.9	50	14	US-10-068-095A-3
9	267	90.8	86	10	US-09-678-380-1
10	267	90.8	86	11	US-09-358-945R-4
11	267	90.8	86	14	US-10-028-410-2
12	267	90.8	86	14	US-10-054-873-4
13	267	90.8	96	10	US-09-947-563-4
14	267	90.8	110	5	US-09-265-854-12
15	267	90.8	110	5	US-09-915-229-5

16	267	90.8	110	10	US-09-804-409A-9	Sequence 9, Appl
17	267	90.8	110	12	US-09-969-748C-6	Sequence 6, Appl
18	267	90.8	110	15	US-10-038-686-1	Sequence 1, Appl
19	267	90.8	110	15	US-10-328-813-2	Sequence 2, Appl
20	267	90.8	117	9	US-09-280-030-63	Sequence 63, Appl
21	267	90.8	130	9	US-09-280-030-62	Sequence 62, Appl
22	266.5	90.6	124	9	US-09-736-611-12	Sequence 12, Appl
23	266.5	90.6	124	9	US-09-740-359-12	Sequence 12, Appl
24	266.5	90.6	124	10	US-09-894-711-12	Sequence 12, Appl
25	266.5	90.6	125	9	US-09-736-611-10	Sequence 10, Appl
26	266.5	90.6	125	9	US-09-740-359-10	Sequence 10, Appl
27	266.5	90.6	125	10	US-09-894-711-10	Sequence 10, Appl
28	266.5	90.6	147	9	US-09-736-611-8	Sequence 8, Appl
29	266.5	90.6	147	9	US-09-740-359-7	Sequence 7, Appl
30	265	90.1	144	9	US-09-736-611-6	Sequence 6, Appl
31	265	90.1	144	9	US-09-740-359-5	Sequence 5, Appl
32	265	90.1	146	10	US-09-894-711-5	Sequence 5, Appl
33	261	88.8	96	10	US-09-947-563-5	Sequence 5, Appl
34	244.5	83.2	144	10	US-09-894-711-7	Sequence 7, Appl
35	219	74.5	46	9	US-09-265-858-132	Sequence 132, App
36	214	72.8	46	9	US-09-265-858-133	Sequence 133, App
37	209	68.0	46	9	US-09-265-858-135	Sequence 135, App
38	195	66.4	46	9	US-09-265-858-134	Sequence 134, App
39	194	66.0	46	9	US-09-205-658-136	Sequence 136, App
40	184	62.6	46	9	US-09-205-658-137	Sequence 137, App
41	183.5	62.4	50	9	US-09-205-658-115	Sequence 115, App
42	173	58.8	32	9	US-09-815-229-14	Sequence 14, Appl
43	173	58.8	32	10	US-09-947-563-6	Sequence 6, Appl
44	173	58.8	56	9	US-09-815-229-10	Sequence 10, Appl
45	173	58.8	156	9	US-09-925-297-805	Sequence 805, App

ALIGNMENTS

RESULT :
US-10-054-873-5
Sequence 5, Application US/10054873
Publication No. US20020164712A1
GENERAL INFORMATION:
APPLICANT: Gan, Zhong Ru
TITLE OF INVENTION: Chimeric Protein Containing an Intramolecular Chaperone-Like Sequence
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew L.P.
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3634
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DCS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/054,873
FILING DATE: 22-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/CN98/G0052
FILING DATE: 31-MAR-1998
APPLICATION NUMBER: US 09/423,100
FILING DATE: 11-DEC-2000
ATTORNEY/AGENT INFORMATION:
NAME: Mycroft, Frank J
REGISTRATION NUMBER: 46,946
REFERENCE/DOCKET NUMBER: 020167-000130US
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 52 amino acids
TYPE: amino acid

```

; STRANDEDNESS: UNKNOWN
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID:
CS-10-054-B73-K

```

```

Query Match      107.0%  Score 294  DB 117  Length 12
Best Local Similarity 100.0%  E-Val 3.5E-20  100%  0%
Matches 52/ Conservative 0/ Miscellaneous 1 100%  0%

QY      2  FYNHILCSHVAEALVLCRGGFFVTKTKGIVDNTCSLSVHENVK 52
Db      1  FYNHILCSHVAEALVLCRGGFFVTKTKGIVDNTCSLSVHENVK 52

```

RESULT 2
US-10-054-073-6
: Sequence 6, Application US/10054873
: Publication No. US2020164712A1

GENERATION: 1
 APPLICANT: Gan, Ziqun Yu
 TITLE OF INVENTION: Chimero Protein Containing an
 Intramolecular Chaperone-like Sequence
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Crow LLP
 STREET: Two Embarcadero Center, Eighth Floor
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94111-3834

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

1 CLASSIFICATION: ~~CONFIDENTIAL~~
 2
 3 PRIORITY APPLICATION DATA:
 4
 5 APPLICATION NUMBER: WO PCT/CN96/00052
 6 FILING DATE: 31-MAR-1998
 7 APPLICATION NUMBER: US 09/423,110
 8 FILING DATE: 11-DEC-2000
 9
 10 ATTORNEY/AGENT INFORMATION:
 11 NAME: MYCROFT, FRANK J.
 12
 13 REGISTRATION NUMBER: 46,946
 14 REFERENCE/DOCKET NUMBER: 020167-00013615
 15
 16 INFORMATION FOR SEQ ID NO: 1:
 17
 18 SEQUENCE CHARACTERISTICS:
 19 LENGTH: 107 amino acids
 20 TYPE: amino acid
 21 STRANDEDNESS: ~~UNKNOWN~~
 22 TOPOLOGY: linear
 23
 24 MOLECULE TYPE: protein
 25
 26 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 27
 28 PS-10-054-873-6

```

Query Match      160.00; Score 254; IP 14; length 17;
Best Local Similarity 100.00; pred. No. 5, 50-29;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps

```

RESULT 3
US-10-054-873-7
; Sequence 7, Application: US/10054873
; Publication No. US20020164712A;
; GENERAL INFORMATION;

APPEZIANI : Guo, Zhong Ku
 TITLE OF INVENTION: Calmaric Protein Containing an
 Intramolecular Chaperone-Like Sequence
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESS: Townsend and Townsend and Crew LLP
 STREET: 100 Embarcadero Center, Eighth Floor
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94111-2634

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatInfo 1.0, Version #1.40
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/1054,673
FILING DATE: 22-Jan-2002
CLASSIFICATION: <Unknown>

PRICE APPLICATION DATA:
APPLICATION NUMBER: WO PCT/CN98/00052
FILING DATE: 31-MAR-1998
APPLICATION NUMBER: US 09/423,100
FILING DATE: 13-DEC-2000
ATTORNEY/AGENT INFORMATION:
NAME: Mycroft, Frank J
REGISTRATION NUMBER: 46,946
REFERENCE/DOCKET NUMBER: 020167-00013005
APPLICATION FOR SEQ ID NO: 7

REFERENCE: DOUGLAS, J. W. J. *PROTEIN*, 1979, 17, 102-110.
 INFORMATION FOR SEQ ID NO: 7:
 LENGTH: 155 amino acids
 TYPE: amino acid
 STRANDEDNESS: <Unknown>
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 7:
 4-873-7

```

Match: 100.0%; Score 294; DB 14; Length 150;
%seq similarity 100.0%; Pred. No. 7.9e-39;
s 52: Conservative 0; Mismatches 0; Indels 0; Gaps
: FVNHCLCGSH:VEALXVLVCGERGFYFPKRGIVEQCTSICSIVOLENYCN 52
|||||
99 FVNHCLCGSH:VEALXVLVCGERGFYFPKRGIVEQCTSICSIVOLENYCN 150
|||||

```

6-935B-5
 Page 5, Application US/69658935B
 Application NO. US20030059177A1
 1. INFORMATION:
 INVENTOR: Dubaque, Yves
 AGENT: Flivartoff, Ellen
 ATTORNEY: Lowman, Henry B.
 TITLE OF INVENTION: METHOD FOR TREATING CARTILAGE DISORDERS
 REFERENCE: P1794R
 PRIOR APPLICATION NUMBER: US/09/856,935B
 FILING DATE: 2002-07-02
 APPLICATION NUMBER: US 60/248,985
 FILING DATE: 2000-11-15
 APPLICATION NUMBER: US 60/204,490
 FILING DATE: 2000-05-16
 NUMBER OF SEQ ID NOS: 153
 NO 5
 PH: 51
 : PR:
 : RISM: Homo sapiens
 6-935B-5

Match	94.7%	Score 278.5;	DB 11;	Length 51;
Local Similarity	98.1%	Pred. No. 2e-27;		

```
Matches 51: Conservative 0: Mismatches 0: Indels 1: Gaps 1:
QY : FVNQHLGGSHLVYALYVCGERGFEYTPKTRGIVEQCCTSCISLYOLENYCN 52
DB : FVNQHLGGSHLVYALYVCGERGFEYTPKTRGIVEQCCTSCISLYOLENYCN 52
DB : FVNQHLGGSHLVYALYVCGERGFEYTPKTRGIVEQCCTSCISLYOLENYCN 51

RESULT 5
US-10-028-410-3
: Sequence 3, Application US/1028410
: Publication No. US20020160955A1
: GENERAL INFORMATION:
: APPLICANT: Debaquie, Yves
: APPLICANT: Lowman, Henry
: TITLE OF INVENTION: PROTEIN VARIANTS
: FILE REFERENCE: P171283-1
: CURRENT APPLICATION NUMBER: US/10/028,410
: CURRENT FILING DATE: 2001-12-19
: PRIOR APPLICATION NUMBER: US/03/477,424
: PRIOR FILING DATE: 2000-01-15
: NUMBER OF SEQ ID NOS: 6
: SEQ ID NO 3
: LENGTH: 51
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-028-410-3

Query Match 94.74: Score 278.5: DB 10: Length 124:
Best Local Similarity 94.34: Pred. No. 1.4e-26:
Matches 51: Conservative 0: Mismatches 0: Indels 1: Gaps 1:
QY : FVNQHLGGSHLVYALYVCGERGFEYTPKTRGIVEQCCTSCISLYOLENYCN 52
DB : FVNQHLGGSHLVYALYVCGERGFEYTPKTRGIVEQCCTSCISLYOLENYCN 52
DB : FVNQHLGGSHLVYALYVCGERGFEYTPKTRGIVEQCCTSCISLYOLENYCN 51

RESULT 6
US-09-894-711-18
: Sequence 18, Application US/09894711
: Patent No. US6020137144A1
: GENERAL INFORMATION:
: APPLICANT: Kjeldsen, Thomas Borglum
: APPLICANT: Ludvigsen, Sverre
: TITLE OF INVENTION: Method for making insulin precursors and
: TITLE OF INVENTION: Insulin precursor analogues having improved formulation
: TITLE OF INVENTION: Yield in yeast
: FILE REFERENCE: 6149,400-US
: CURRENT APPLICATION NUMBER: US/09/894,711
: CURRENT FILING DATE: 2001-06-28
: PRIOR APPLICATION NUMBER: PA 2000 00443
: PRIOR FILING DATE: 2000-03-17
: PRIOR APPLICATION NUMBER: PA 1999 01869
: PRIOR FILING DATE: 1999-12-29
: PRIOR APPLICATION NUMBER: 63/211,081
: PRIOR FILING DATE: 2000-06-13
: PRIOR APPLICATION NUMBER: 63/181,450
: PRIOR FILING DATE: 2000-02-10
: PRIOR APPLICATION NUMBER: 63/743,359
: PRIOR FILING DATE: 2000-12-19
: NUMBER OF SEQ ID NOS: 20
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 18
: LENGTH: 124
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Synthetic
US-09-894-711-18

Query Match 93.74: Score 275.5: DB 10: Length 124:
Best Local Similarity 94.34: Pred. No. 1.4e-26:
Matches 50: Conservative 1: Mismatches 1: Indels 1: Gaps 1:
QY : FVNQHLGGSHLVYALYVCGERGFEYTPKTRGIVEQCCTSCISLYOLENYCN 52
DB : FVNQHLGGSHLVYALYVCGERGFEYTPKTRGIVEQCCTSCISLYOLENYCN 138
DB : FVNQHLGGSHLVYALYVCGERGFEYTPKTRGIVEQCCTSCISLYOLENYCN 138

RESULT 8
US-10-056-009A-3
: Sequence 3, Application US/10066009A
: Publication No. US20020165155A1
: GENERAL INFORMATION:
: APPLICANT: Schaffer, Michelle
: APPLICANT: Jitsch, Mark
: APPLICANT: Vajdos, Felix
: TITLE OF INVENTION: CRYSTALLIZATION OF IGF-1
: FILE REFERENCE: P1669R1
```

```
QY 1 FVNQHLGGSHLVYALYVCGERGFEYTPKTRGIVEQCCTSCISLYOLENYCN 52
DB 72 FVNQHLGGSHLVYALYVCGERGFEYTPKTRGIVEQCCTSCISLYOLENYCN 124

RESULT 7
US-09-861-687-19
: Sequence 19, Application US/C9861687
: Publication No. US20020193292A1
: GENERAL INFORMATION:
: APPLICANT: Markussen, Jan
: APPLICANT: Jonassen, Ib
: APPLICANT: Havelund, Sverre
: APPLICANT: Brandt, Jakob
: APPLICANT: Kurtzals, Peter
: APPLICANT: Hansen, Hertz Per
: APPLICANT: Kaarsholm, Niels Christian
: TITLE OF INVENTION: INSULIN DERIVATIVES
: NUMBER OF SEQUENCES: 26
: CORRESPONDENCE ADDRESS:
: ADDRESS: NO. US20020193292A1o No. US20020193292A1disk of No. US20
: STREET: 405 Lexington Avenue, 64th Floor
: CITY: New York
: STATE: New York
: COUNTRY: United States of America
: ZIP: 10174-6401
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/861,687
: FILING DATE: 21-May-2001
: CLASSIFICATION: <unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 68/932,082
: FILING DATE: 16-DEC-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Lambiris, Elias J.
: REGISTRATION NUMBER: 33,728
: REFERENCE/DOCKET NUMBER: 4341,204-US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-867-0123
: TELEFAX: 212-878-9655
: INFORMATION FOR SEQ ID NO: 19:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 138 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-861-687-19

Query Match 93.74: Score 275.5: DB 10: Length 138:
Best Local Similarity 94.34: Pred. No. 1.4e-26:
Matches 50: Conservative 1: Mismatches 1: Indels 1: Gaps 1:
QY 1 FVNQHLGGSHLVYALYVCGERGFEYTPKTRGIVEQCCTSCISLYOLENYCN 52
DB 86 FVNQHLGGSHLVYALYVCGERGFEYTPKTRGIVEQCCTSCISLYOLENYCN 138
DB 86 FVNQHLGGSHLVYALYVCGERGFEYTPKTRGIVEQCCTSCISLYOLENYCN 138

RESULT 9
US-10-056-009A-3
: Sequence 3, Application US/10066009A
: Publication No. US20020165155A1
: GENERAL INFORMATION:
: APPLICANT: Schaffer, Michelle
: APPLICANT: Jitsch, Mark
: APPLICANT: Vajdos, Felix
: TITLE OF INVENTION: CRYSTALLIZATION OF IGF-1
: FILE REFERENCE: P1669R1
```



```
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/10/054,873
? FILING DATE: 22-Jan-2002
? CLASSIFICATION: <Unknown>
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: WO PCT/CN98/00052
? FILING DATE: 31-Mar-1998
? APPLICATION NUMBER: US 09/423,100
? FILING DATE: 11-DEC-2003
? ATTORNEY/AGENT INFORMATION:
? NAME: MYCROFT, Frank J
? REGISTRATION NUMBER: 46,946
? REFERENCE/DOCKET NUMBER: 020167-000130/5
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 86 amino acids
? TYPE: amino acid
? STRANDEDNESS: <Unknown>
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-054-873-4

Query Match 90.8% Score 267; DB 14; Length 86;
Best Local Similarity 60.5%; Pred. No. 1.1e-25;
Matches 52; Conservative 0; Mismatches 0; Indels 54; Gaps 1;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKT----- 86
DB 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAFLQVGQVHLGGPSGASLQPLALEG 68
QY 31 ----RGIVEQCCTSIICSLYLENYCN 52
DB 61 SLQKRGIVEQCCTSIICSLYLENYCN 86

RESULT 13
US-09-947-563-4
? Sequence 4, Application US/09/47563
? Patent No. US20020156234A1
? GENERAL INFORMATION:
? APPLICANT: Subroder, Franz-Josef;
? Keller, Reinhold
? TITLE OF INVENTION: Improved process for obtaining
? insulin precursors having correctly bonded cysteine bridges
? NUMBER OF SEQUENCES: 7
? CORRESPONDENCE ADDRESS
? ADDRESSEE: Finckhan, Henderson, Farrabow, Garrett &
? Dunner
? STREET: 1400 I Street, N.W.
? CITY: Washington
? STATE: D.C.
? COUNTRY: USA
? ZIP: 20005-3315
? COMPUTER READABLE FORM
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/947,563
? FILING DATE: 07-Sep-2001
? CLASSIFICATION: <Unknown>
? PRIOR APPLICATION DATA
? APPLICATION NUMBER: 09/134,836
? FILING DATE: <Unknown>
? ATTORNEY/AGENT INFORMATION:
? NAME: Leslie McDowell
? REGISTRATION NUMBER: 34,872
? REFERENCE/DOCKET NUMBER: 02481.1500-00000
```

```
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (202) 408-4000
? TELEFAX: (202) 408-4400
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 96 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? ORIGINAL SOURCE:
? ORGANISM: Escherichia coli
? FEATURE:
? NAME/KEY: Protein
? LOCATION: 1..96
? SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-947-563-4

Query Match 90.8% Score 267; DB 9; Length 96;
Best Local Similarity 60.5%; Pred. No. 1.1e-25;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKT----- 30
DB 21 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAFLQVGQVHLGGPSGASLQPLALEG 70
QY 31 ----RGIVEQCCTSIICSLYLENYCN 52
DB 71 SLQKRGIVEQCCTSIICSLYLENYCN 96

RESULT 14
US-09-205-658-125
? Sequence 125, Application US/09205658
? Patent No. US20010029617A:
? GENERAL INFORMATION:
? APPLICANT: Ruykun, Gary
? APPLICANT: Ogg, Scott
? TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
? TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
? FILE REFERENCE: 00786/351004
? CURRENT APPLICATION NUMBER: US/09/205,658
? CURRENT FILING DATE: 1998-12-03
? EARLIER APPLICATION NUMBER: 08/857,076
? EARLIER FILING DATE: 1997-05-15
? EARLIER APPLICATION NUMBER: 08/888,534
? EARLIER FILING DATE: 1997-07-07
? EARLIER APPLICATION NUMBER: US98/10080
? EARLIER FILING DATE: 1998-05-15
? NUMBER OF SEQ ID NOS: 328
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 125
? LENGTH: 110
? TYPE: PRT
? ORGANISM: Homo sapiens
? US-09-205-658-125

Query Match 90.8% Score 267; DB 9; Length 110;
Best Local Similarity 60.5%; Pred. No. 1.2e-25;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 1 FVNQHLGSHLVEALYLVCGERGFFYTPKT----- 30
DB 25 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAFLQVGQVHLGGPSGASLQPLALEG 84
QY 31 ----RGIVEQCCTSIICSLYLENYCN 52
DB 85 SLQKRGIVEQCCTSIICSLYLENYCN 110

RESULT 15
US-09-615-229-3
? Sequence 3, Application US/09815229
```

```

: Patent No. US2002005862A1
: GENERAL INFORMATION:
: APPLICANT: Filvaroff, Elton H.
: APPLICANT: Okuno, Franklin W.
: TITLE OF INVENTION: USE OF INSULIN FOR THE TREATMENT OF CARCINOGENESIS AND SURVIVAL
: FILE REFERENCE: P1786R1US
: CURRENT APPLICATION NUMBER: US 60/815,229
: CURRENT FILING DATE: 2001-03-22
: PRIOR APPLICATION NUMBER: US 60/52,103
: PRIOR FILING DATE: 2000-03-24
: NUMBER OF SEQ ID NOS: 17
: SEQ ID NO 3
: LENGTH: 116
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-815-229-3

```

```

Query Match      90.88; Score 257; Db 9; Length 116;
Best Local Similarity 60.58; Pred. No. 1,26-26;
Matches 52; Conservative 0; Mismatches 9; Gaps 14
QY      1  FVNQLCGSHLVEALYLGVGGGFFYTKT----- -- 30
Db      25  FVNQLCGSHLVEALYLGVGGGFFYTKT----- -- 30
QY      31  ----RGIVEQCCTSTCSLFOLENYCN 52
Db      85  SLCRGIVEQCCTSTCSLFOLENYCN 110

```

```

Search completed: September 15, 2003, 12:23:36
Job time : 22.8065 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Inc.

OM protein - protein search, using sw model

Run on: September 15, 2003, 11:56:45 : Search time 8.67348 seconds
(without alignments)
583.284 Million cell updates/sec

Title: US-09-423-100-5

Perfect score: 294

Sequence: 1 FVNQHLCGSHLVEA...IVEQCCTICSLEYOLENYCN 52

Scoring table: BLOSUM62

Gap: 10.0, Gapext: 0.5

Searched: 283308 seqs, 96162682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match: 0%

Maximum Match: 100%

Listing first 45 summaries

Database : PIR_76:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by change to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	273.5	93.0	51	INMSP	insulin - sperm whale
2	273.5	93.0	51	INMSE	insulin - finback whale
3	273.5	93.0	51	INMSE	insulin - elephant
4	273.5	92.9	96	PC7582	epidermal growth factor
5	271.5	92.3	51	INMSE	insulin - hamster
6	268.5	91.3	51	INMSSP	insulin - Egyptian
7	267.5	91.0	51	A59151	insulin precursor
8	267	90.8	110	IPHC	insulin precursor
9	267	90.8	110	B42179	insulin precursor
10	267	90.8	110	A42179	insulin precursor
11	267	90.8	110	CQ1178	insulin precursor
12	263.5	89.6	51	INMHS	insulin - seal whale
13	263.5	89.6	51	INMHS	insulin - goat
14	263.5	89.6	51	INMKA	insulin - Arabian
15	263	89.5	84	IPGC	insulin precursor
16	263	89.5	110	INMSE	insulin precursor
17	262.5	89.3	51	INMSE	insulin - rat
18	262	89.1	110	IPGC	insulin precursor
19	261.5	88.9	51	INMKSQ	insulin - common
20	260	88.4	110	I48166	insulin precursor
21	258.5	87.9	135	IPBO	insulin precursor
22	256.5	87.2	51	JQ0362	insulin precursor
23	252.5	85.9	77	INSH	insulin - North Am
24	252	85.7	86	IPHC	insulin precursor
25	251.5	85.5	51	INMSE	insulin - chick
26	251	85.4	128	A39883	insulin precursor
27	250	85.0	108	INMSE	insulin precursor
28	249	84.7	110	IPRT1	insulin precursor
29	248.5	84.5	51	INMSE	insulin - dog

insulin 2 precursor
insulin 2 precursor
insulin 11 - North
insulin 12 - North
insulin - crested
insulin - turkey
insulin - ostrich
insulin - black-be
insulin - slider t
insulin precursor
insulin - Amphiuma
insulin - duckbill
insulin precursor
insulin precursor
insulin - American
insulin - alligator

ALIGNMENTS

RESULT 1

INMSE

insulin - sperm whale

C.Species: Physeter catodon (sperm whale)

C.Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999

C.Accession: A93142; A90082

R.Ishihara, Y.; Saito, T.; Ito, Y.; Fujino, M.

Nature 161, 1466-1469, 1958

Title: Structure of sperm- and seal-whale insulins and their breakdown by whale I

A.Reference number: A93142

A.Accession: A93142

A.Molecule type: protein

A.Residues: 1-30;31-51<1SH>

R.Ishihara, Y.; Saito, T.; Ito, Y.; Naughton, M.A.

Arch. Biochem. Biophys. 85, 427-428, 1956

Title: Species differences in insulin.

A.Reference number: A90082

A.Accession: A90082

A.Molecule type: protein

A.Residues: 1-30;31-51<HAR>

C.Species: insulin

C.Keywords: hormone; pancreas

F.1-30/Domain: insulin chain B #status experimental <BCH>

F.1-30,31-51/Product: insulin #status experimental <MAT>

F.1-51/Domain: insulin chain A #status experimental <ACH>

F.1-37,19-50,36-42/Disulfide bonds: #status predicted

Query Match 93.0% Score 273.5; DB 1; Length 51;

Best local similarity 96.2%; Pred. No. 1.5e-24;

Matches 50; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY : FVNQHLCGSHLVEA...YVCGERGFFYPTKRGIVEQCCTICSLEYOLENYCN 52

DD : 1 FVNQHLCGSHLVEA...YVCGERGFFYPTKA-GIVEQCCTICSLEYOLENYCN 51

RESULT 2

INMSE

insulin - finback whale (tentative sequence)

C.Species: Balenoptera physalus (finback whale, common rorqual)

C.Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999

C.Accession: A91918

R.Hama, H.; Titani, K.; Sakaki, S.; Narita, K.

J. Biochem. 56, 265-293, 1964

Title: The amino acid sequence in fin-whale insulin.

A.Reference number: A91918

A.Accession: A91918

A.Molecule type: protein

A.Residues: 1-30;31-51<HAR>

C.Species: insulin

C.Keywords: hormone; pancreas

F.1-30/Domain: insulin chain B #status experimental <BCH>


```

A:Accession: B42179
A:Molecule type: DNA
A:Residues: 1-110 <SEI>
A:Cross-references: EMBL:X61092; NID:922808; PID:CAA1435.1; PID:922909
A:Note: sequence extracted from NCB1 backbone (NID:95.95; NID:95.14)
A:Peterson, J.D.; Mehrlich, S.; Gyer, P.E.; Steneger, D.F.
J. Biol. Chem. 247, 4865-4871, 1972
A:Title: Determination of the amino acid sequence of the monkey, sheep, and bat proinsulin
A:Reference number: A92111; MUI:72258016; PMID:4624369
A:Accession: A05242
A:Molecule type: protein
A:Residues: 57-87 <PPT>
C:Genetics:
A:Introns: 63/
A:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-54/Domain: insulin chain B #status predicted <ACH>
F:55-87/Domain: insulin chain A #status predicted <MAP>
F:57-87/Domain: connecting peptide #status experimental <CHP>
F:90-110/Domain: insulin chain A #status predicted <CH>
F:91-96.43-109.95-100/Disulfide bonds: #status predicted
Query Match 90.8%; Score 267; DB 2; Length 110;
Best Local Similarity 60.5%; Pred. No. 1.6e-23;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1.
QY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKT----- 45
DB 25 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDQVQVELGGGPGAGSIQPLALEG 4
QY 31 ----RGVEQCCTSCISLYQLENYCN 52
DB 85 SLQKRGIVEGCTSCISLYQLENYCN 110

RESULT 10
A42179
insulin precursor - chimpanzee
C:Species: Pan troglodytes (chimpanzee)
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 16-Jul-1999
C:Accession: A42179; S22058
R:Seino, S.; Bell, G.I.; Li, W.H.
Mol. Biol. Evol. 9, 193-203, 1992
A:Title: Sequences of primate insulin genes support the hypothesis of a slower rate of
A:Reference number: A42179; MUI:92219553; PMID:1560757
A:Accession: A42179
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-110 <SEI>
A:Cross-references: EMBL:X61092; NID:938251; PID:CAA1435.1; PID:922909
A:Note: sequence extracted from NCB1 backbone (NID:95.95)
C:Genetics:
A:Introns: 63/
A:Superfamily: insulin

Query Match 90.8%; Score 267; DB 2; Length 110;
Best Local Similarity 60.5%; Pred. No. 1.6e-23;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1.
QY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKT----- 45
DB 25 FVNHLCGSHLVEALYLVCGERGFFYTPKTRREAEDQVQVELGGGPGAGSIQPLALEG 43
QY 31 ----RGVEQCCTSCISLYQLENYCN 52
DB 85 SLQKRGIVEGCTSCISLYQLENYCN 110

RESULT 11
J00178
insulin precursor - crab-eating macaque
C:Species: Macaca fascicularis (crab-eating macaque)

```

```

C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 16-Jul-1999
C:Accession: J00178
C:Species: Capra aegagrus hircus (domestic goat)
C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 15-Jul-1999
C:Accession: A01586
R:Smith, L.F.
Am. J. Med. 40, 662-666, 1966
A:Title: Species variation in the amino acid sequence of insulin.
A:Reference number: A90029; MUI:66160119; PMID:5949593
A:Accession: A01586
A:Molecule type: protein
A:Residues: 1-30;31-51 <SMI>

Query Match 89.6%; Score 263.5; DB 3; Length 51;
Best Local Similarity 92.3%; Pred. No. 2.1e-23;
Matches 48; Conservative 0; Mismatches 3; Indels 1; Gaps 1.
QY 1 FVNHLCGSHLVEALYLVCGERGFFYTPKTRGIVEGCTSCISLYQLENYCN 52
DB 3 FVNHLCGSHLVEALYLVCGERGFFYTPKA-GIVEGCTSCISLYQLENYCN 51

RESULT 14
J00178
insulin - goat
C:Species: Capra aegagrus hircus (domestic goat)
C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 15-Jul-1999
C:Accession: A01586
R:Smith, L.F.
Am. J. Med. 40, 662-666, 1966
A:Title: Species variation in the amino acid sequence of insulin.
A:Reference number: A90029; MUI:66160119; PMID:5949593
A:Accession: A01586
A:Molecule type: protein
A:Residues: 1-30;31-51 <SMI>

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Computer Ltd.

OM protein - protein search, using sw mode

Run on: September 15, 2003, 11:54:00 : Search time 4.81588 seconds
(without alignment)
504.033 Million cell updates/sec

Title: US-09-423-100-5
Perfect score: 234
Sequence: 1 FVQHICGSHLVEALYLVCG.....IVEGCGSICSLYLENYCN 52

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127553 seqs, 47025705 residues

Total number of hits satisfying chosen parameters: 127553

Minimum DR seq length: 0
Maximum DR seq length: 230000000

Post-processing: Minimum Match 1%

Listing first 15 summaries

Database : SwissProt_41*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DR	ID	Description
1	273.5	93.0	51	1	INS_BALPH	1	INS BALPH
2	271.5	91.0	51	1	INS_ELEMA	1	INS ELEMA
3	268.5	91.3	51	1	INS_ACOCA	1	INS ACOCA
4	267.0	90.4	110	1	INS_GERAR	1	INS GERAR
5	267.0	90.8	110	1	INS_BOVAN	1	INS BOVAN
6	267.0	90.8	110	1	INS_MACFA	1	INS MACFA
7	267.0	90.8	110	1	INS_PANTA	1	INS PANTA
8	263.5	89.6	51	1	INS_RALBO	1	INS RALBO
9	263.5	89.6	51	1	INS_GAMOR	1	INS GAMOR
10	263.5	89.6	51	1	INS_GAPHI	1	INS GAPHI
11	263.0	89.5	108	1	INS_PEG	1	INS PEG
12	263.0	89.5	110	1	INS_PABIT	1	INS PABIT
13	263.0	89.5	110	1	INS_SPTIR	1	INS SPTIR
14	262.5	89.3	51	1	INS_FELCA	1	INS FELCA
15	262.0	89.1	110	1	INS_CANEA	1	INS CANEA
16	260.0	88.4	110	1	INS_ORILO	1	INS ORILO
17	258.5	87.9	105	1	INS_BOVIN	1	INS BOVIN
18	257.0	87.4	110	1	INS_PSAOR	1	INS PSAOR
19	256.5	87.2	51	1	INS_DIDKA	1	INS DIDKA
20	254.5	86.6	105	1	INS_SHEEP	1	INS SHEEP
21	252.0	85.7	86	1	INS_FORSE	1	INS FORSE
22	251.5	85.5	51	1	INS_CHIBR	1	INS CHIBR
23	251.0	85.4	108	1	INS_AOYFR	1	INS AOYFR
24	250.0	85.0	108	1	INS_MOUSE	1	INS MOUSE
25	249.0	84.7	110	1	INS_LBAT	1	INS LBAT
26	248.5	84.5	51	1	INS_RANAN	1	INS RANAN
27	248.0	84.4	110	1	INS2_MOUSE	1	INS2_MOUSE
28	248.0	84.4	110	1	INS2_RAT	1	INS2_RAT
29	246.0	83.7	52	1	INS_AC1G0	1	INS AC1G0
30	244.5	83.2	51	1	INS_HYSCR	1	INS HYSCR
31	244.5	83.2	51	1	INS_TRASC	1	INS TRASC
32	238.5	81.1	103	1	INS_SELRF	1	INS SELRF
33	235.5	80.1	51	1	INS_ORNAN	1	INS ORNAN

P01332 gallus gal.
P01333 anas platyr
P01703 alligator m
P09476 lepisosteus
P12708 zaocys dhum
P01334 crocatus at
P12706 xenopus lae
P12707 xenopus lae
P14806 petromyzon
P09477 platichthys
P29335 amia calva
P23187 oncornynchu

ALIGNMENTS

RESULT 1
TE INS_BALPH STANDARD; PRT: 51 AA.
AC P01312;
DE 21-JUN-1985 (Rel. 01, Created)
DC 21-JUN-1985 (Rel. 01, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
ID Insulin.
IN INS.
CS Balanoptera physalius (Finback whale) (Common rorqual), and
US Physeter catodon (Sperm whale) (Physeter macrocephalus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;
CC Balanopteridae; Balanoptera.
CX NCBI_FaxID:9770, 9755;
FN 111
RP PARTIAL SEQUENCE.
RC SPECIES=B.physalius;
EA Hama H., Titani K., Sakaki S., Narita K.;
PT "The amino acid sequence in fin-whale insulin.";
EL J. Biochem. 56:285-291(1964).
RN 12
RQ SEQUENCE.
RA Ishihara Y., Saito J., Ito Y., Fujino M.;
PT "Structure of sperm- and sei-whale insulins and their breakdown by
KL whale pepsin.";
RL Nature 181:1448-1453(1958).
RN 11
RQ SEQUENCE.
RC SPECIES=P.catodon;
RA Harris J., Sanger F., Naughton M.A.;
PT "Species differences in insulin.";
EL Arch. Biochem. Biophys. 65:427-439(1956).
CC -1- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
DISULFIDE BONDS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR PIR: A91918; INWHF.
DR PIR: A93142; INWHF.
DR RSP: P01317; IAPH.
DR InterPro: IPR004825; Ins/IGF/relax.
DR SMART: SM00078; IIGF; 1
DR PROSITE: PS00262; INSULIN; 1.
KW Insulin family; Hormone; Glucose metabolism.
FT CHAIN 1 30 INSULIN B CHAIN.
FT NGL_CONS 30 31
FT CHAIN 31 51 INSULIN A CHAIN.
FT DISULFID 7 37 INTERCHAIN.
FT DISULFID 19 50 INTERCHAIN.
FT DISULFID 36 41
SQ SEQUENCE 51 AA: 5766 MW: 5007B514691A7CDD CRC64:

```

GX NCBI_Taxid=10968;
GN [1]
KN COMPOSITION.
KP MEDLINE=72258454; PubMed=5028210;
RA Bueweli H.F., Humbel R.E.;
RC "Isolation and partial structural analysis of insulin from mouse (Mus musculus) and spiny mouse (Acomys kuhirous).";
RG Poppe-Reyler's Z. Physiol. Chem. 353:444-450(1972);
CC -I- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -II- SYBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO DISULFIDE BONDS.
CC -III- SUBCELLULAR LOCATION: Secreted.
CC -IV- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC FIR: A01591; INMSEP.
CC HSSE: P01308; ILYN.
CC ZlotarP/O: IPROCA825; Ins/IGF/relax.
CC SWAR: SMOO07H; IGF; J.
CC PROSITE: PS00262; INSULIN; 1.
CC Insulin family; Hormone; Glucose metabolist.
FT CRAN 1 30 INSULIN B CHAIN.
FT GRNCONS 30 3;
FT GRAIN 3 3;
FT DISULFID 7 37 INTERCHAIN (BY SIMILARITY).
FT DISULFID 19 50 INTERCHAIN (BY SIMILARITY).
FT DISULFID 36 41 BY SIMILARITY.
SQ SEQUENCE 51 AA: 5766 MW: 992H2D8E629047D30 CRC64;

Cory Hatch          91.3%; Score 268.5; DB 1; Length 51;
Best Local Similarity 92.3%; Pred.No. 2e-26;
Matches 48; Conservative 3; Mismatches 0; Indels 1; Gaps 1

CY 1 FYVHQLCSSHLVEALVLCGRGFFTPKPGIVEGGCTSCISLYOLENYCN 52
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1 FVHQLCSSHLVEALVLCGRGFFTPKS-GIVDQCISCSLYOLENYCN 51

RESULT 4
INS_CERAE
LD INS_CERAE STANDARD; PRT; 110 AA.
AC P30407; P01309;
ET 01-APR-1993 (Rel. 25, Created)
LT 01-APR-1993 (Rel. 25, Last sequence update)
ST 01-FEB-1996 (Rel. 33, Last annotation update)
DE Insulin precursor.
GN INS.
CC Cercopithecus aethiops (Green monkey) (Grivet).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae.
CC Cercopithecoinae; Cercopithecus.
CC NCBI_Taxid=9534;
GN [1]
KN SEQUENCE FROM N.A.
RX MEDLINE=92219953; PubMed=1560757;
RA Seino S., Bell G.L., Li W.;
CC "Sequences of primate insulin genes support the hypothesis of a slower rate of molecular evolution in humans and apes than in monkeys.";
CC Mol. Biol. Evol. 9:133-203(1992).
RN [2];
RP SEQUENCE OF 57-87.
RX MEDLINE=72258016; PubMed=4626369;
PA Peterson S.D., Nefflich S., Oyer P.E., Steiner D.F.;
KT "Determination of the amino acid sequence of the monkey, sheep, and dog proinsulin C-peptides by a semi-micro Raman degradation procedure."
KL J. Biol. Chem. 247:4866-4871(1972).
CC -I- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
```


RP SYNTHESIS OF 57-57.
RX MEDLINE=75540007; PubMed=4923504;
RA Naichan V.K.;
RT "Studies on polypeptides. IV. The synthesis of C peptide of human
RT proinsulin";
RL Hoppe-Seyler's Z. Physiol. Chem. 354:155-162(1973).
RN [13]
RP SYNTHESIS OF 65-69 AND 70-73.
RX MEDLINE=73161263; PubMed=4598559;
RA Geiser R., Volk A.;
RT "Synthesis of peptides with the properties of human proinsulin C
RT peptides (hc peptide). 3. Synthesis of the sequences 55-71 and 6-14
RT of human proinsulin C peptides";
RL Chem. Ber. 106:199-205(1973).
RN [14]
RP SYNTHESIS OF 84-87.
RX MEDLINE=73161261; PubMed=4598553;
RA Geiser R., Jaeger G., Keonig W., Treuth G.;
RT "Synthesis of peptides with the properties of human proinsulin C
RT peptides (hc peptide). I. Scheme for the synthesis and preparation of
RT the sequence 28-31 of human proinsulin C peptide";
RL Chem. Ber. 106:188-192(1973).
RN [15]
RP VARIANTS LOS ANGELES SER-4b.
RX MEDLINE=84316053; PubMed=6212455;
RA Haceda M., Chao S., Kwok S.C.M., Rubenstein A.B., Steiner D.F.;
RT "Studies on mutant human insulin genes: identification and sequence
RT analysis of a gene encoding (SerB34, Insulin)";
RL Proc. Natl. Acad. Sci. U.S.A. 50:6366-6370(1983).
RN [16]
RP VARIANTS LOS ANGELES SER-4b AND CHICAGO LEU-4b.
RX MEDLINE=84170233; PubMed=6424111;
RA Shoelson S., Fickova M., Haneda M., Nahum A., Musso G., Kaiser E.F.;
RT "Identification of a mutant human insulin predicted to contain a
RT serine-for-phenylalanine substitution";
RL Proc. Natl. Acad. Sci. U.S.A. 80:7390-7394(1983).
RN [17]
RP VARIANT PROVIDENCE ASP-34.
RX MEDLINE=87175640; PubMed=3470784;
RA Chan S.J., Seino S., Gruppiso P.A., Schwartz R., Steiner D.F.;
RT "A mutation in the B chain coding region is associated with impaired
RT proinsulin conversion in a family with hyperproinsulinemia";
RL Proc. Natl. Acad. Sci. U.S.A. 84:2194-2197(1987).
RN [18]
RP VARIANT WAKAYAMA LEU-52.
RX MEDLINE=87058122; PubMed=3337011;
RA Sakura H., Tamoto Y., Sakamoto Y., Kuzuya T., Hirata H.;
RT "Structurally abnormal insulin A5 (Val¹ Glu) isolated from the pancreas";
RL J. Clin. Invest. 78:1666-1672(1986).
RN [19]
RP VARIANT HIS-89.
RX MEDLINE=90317621; PubMed=2196279;
RA Harbetti F., Raben N., Kadowaki T., Cama A., Arcill E., Gabay K.B.;
RT "Two unrelated patients with familial hyperproinsulinemia due to a
RT mutation substituting histidine for arginine at position 89 in the
RT proinsulin molecule: identification of the mutation by direct
RT sequencing of genomic deoxyribonucleic acid amplified by polymerase
RT chain reaction";
RL J. Clin. Endocrinol. Metab. 71:164-169(1990).
RN [20]
RP VARIANT HIS-89.
RX MEDLINE=85261994; PubMed=4019786;
RA Shibasaki Y., Kawakami T., Kanazawa Y., Akazawa Y., Takaya F.;
RT "Posttranslational cleavage of proinsulin is blocked by a point
RT mutation in familial hyperproinsulinemia";
RL J. Clin. Invest. 76:378-380(1985).
RN [21]
RP VARIANT KYOTO LEU-83.
RX MEDLINE=92291307; PubMed=1401997;
RA Yano H., Kitano N., Morimoto M., Polonsky R.S., Iwata H., Seino Y.;
RT "A novel point mutation in the human insulin gene giving rise to
RT hyperproinsulinemia (proinsulin Kyoto)";
RL J. Clin. Invest. 89:1902-1907(1992).
RN [22]
RP STRUCTURE BY NMR.
RX MEDLINE=91104966; PubMed=2271664;
RA Hua Q.-X., Weiss M.A.;
RT "Toward the solution structure of human insulin: sequential 2D 1H NMR
RT assignment of a des-pentapeptide analogue and comparison with crystal
RT structure";
RL Biochemistry 29:10545-10555(1996).
RN [23]
RP STRUCTURE BY NMR.
RX MEDLINE=91242467; PubMed=2036420;
RA Hua Q.-X., Weiss M.A.;
RT "Comparative 2D NMR studies of human insulin and des-pentapeptide
RT insulin: sequential resonance assignment and implications for protein
RT dynamics and receptor recognition";
RL Biochemistry 30:5505-5515(1991).
RN [24]
RP STRUCTURE BY NMR.
RX MEDLINE=91265527; PubMed=1646635;
RA Hua Q.-X., Weiss M.A.;
RT "Two-dimensional NMR studies of Des-(B26-B30)-insulin: sequence-
RT specific resonance assignments and effects of solvent composition";
RL Biochim. Biophys. Acta 1078:101-110(1991).
RN [25]
Query Match 90.8%; Score 267; DB 1; Length 110;
Best Local Similarity 60.5%; Pred. No. 6.3e-26;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;
QY 1 FVNQCGSLVEALVVGGERGFYTPKT----- 30
DB 25 FVNQCGSLVEALVVGGERGFYTPKTREAEQLVGVELGGGCGAGSLOPLALEG 84
QY 31 ----RGIVECCCTSCSLYQLENYCN 52
DB 85 SLQKRGIVECCCTSCSLYQLENYCN 110
RESULT 6
INS_MACFA
ID INS_MACFA STANDARD; PRT: 110 AA.
AC P30406; P01309;
DT 21-JUL-1986 (Rel. 01, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
LT 01-FEB-1996 (Rel. 33, Last annotation update)
FE Insulin precursor.
GN INS
CS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
CC Cercopitheciinae; Macaca.
CC NCBI_TaxID=9541;
FN [1]
PP SEQUENCE FROM N.A.
RX MEDLINE=850303474; PubMed=6184262;
RA Wetkham W., Gronenberg C., Heineweb M., Wengenmayer F.,
RA Winnacker E.-L.;
RT "The nucleotide sequence of cDNA coding for preproinsulin from the
RT primate Macaca fascicularis";
RL Gene 19:179-183(1982).
RN [2]
CC -1- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -1- CELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE INSULIN/TGF/RELAXIN FAMILY.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/about-us/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: J00336; AAA36849.1; -
 DR PIR: Q01178; JQ0178.
 DR HSSP: P01408; IAT0.
 DR InterPro: IPR004825; Ins/IGF/relax.
 DR Pfam: PF00049; Insulin; 1.
 DR PRINTS: PR00277; INSULIN.
 DR SMART: SM00078; IIGF; 1.
 DR PROSITE: PS00262; INSULIN; 1.
 KW Insulin family; Hormone; Glucose metabolism; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 54 INSULIN B CHAIN.
 FT PROPEP 57 87 C PEPTIDE.
 FT CHAIN 90 110 INSULIN A CHAIN.
 FT DISULFID 91 96 INTERCHAIN.
 FT DISULFID 43 109 INTERCHAIN.
 FT DISULFID 95 100
 SQ SEQUENCE 110 AA: 1199; MW: 836633AP0A420F9 C6C64;

Query Match 90.8% Score 267; DR 1; Length 110;
 Best Local Similarity 60.5% Pred. No. 6.3e-26;
 Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 1 FVNHGSGHSELVYALVCGERGFFYTPKT----- 30
 DB 25 FVNHGSGHSELVYALVCGERGFFYTPKTRAEADLVQGVFVGSGPGAGSLQPLAEG 84
 QY 31 ----RGIVEGCGCTSGSLYLENYCN 52
 DB 85 SLGKPGIVEGCGCTSGSLYLENYCN 110

RESULT 7
 INS_PANTR
 ID INS_PANTR STANDARD; PRT; 110 AA.
 AC P30410;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Insulin precursor.
 GN INS.
 OS Pan troglodytes (Chimpanzee).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 CX NCBI_TaxID=9598;
 RN [1]
 RP MEDLINE:92219953; PubMed:1560757;
 RA Seino S., Bell G.I., Li W.
 RT "Sequences of primate insulin genes support the hypothesis of a
 RT slower rate of molecular evolution in humans and apes than in
 RT monkeys."
 RL Mol. Biol. Evol. 9:193-203(1992).
 CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
 CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
 CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
 CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
 CC -!- SURUNIT: HETEROIDIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
 CC DISULFIDE BONDS.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/about-us/>
 CC or send an email to license@isb-sib.ch).

CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: X61069; CAA43403.1; -
 DR PIR: A42179; A42179.
 DR DB: IIGF; 29-MAR-00.
 DR InterPro: IPR004825; Ins/IGF/relax.
 DR Pfam: PF00049; Insulin; 1.
 DR PRINTS: PR00277; INSULIN.
 DR SMART: SM00078; IIGF; 1.
 DR PROSITE: PS00262; INSULIN; 1.
 KW Insulin family; Hormone; Glucose metabolism; Signal; 3D-structure.
 FT SIGNAL 1 24
 FT CHAIN 25 54 INSULIN B CHAIN.
 FT PROPEP 57 87 C PEPTIDE.
 FT CHAIN 90 110 INSULIN A CHAIN.
 FT DISULFID 91 96 INTERCHAIN.
 FT DISULFID 43 109 INTERCHAIN.
 FT DISULFID 95 100
 SQ SEQUENCE 110 AA: 12025 MW: 41E8DF79837CEEF5 CRC64;

Query Match 90.8% Score 267; DR 1; Length 110;
 Best Local Similarity 60.5% Pred. No. 6.3e-26;
 Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 1 FVNHGSGHSELVYALVCGERGFFYTPKT----- 30
 DB 25 FVNHGSGHSELVYALVCGERGFFYTPKTRAEADLVQGVFVGSGPGAGSLQPLAEG 84
 QY 31 ----RGIVEGCGCTSGSLYLENYCN 52
 DB 85 SLGKPGIVEGCGCTSGSLYLENYCN 110

RESULT 8
 INS_BALBO
 ID INS_BALBO STANDARD; PRT; 51 AA.
 AC P3134;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Insulin.
 GN INS.
 OS Balaeoptera borealis (Sei whale).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;
 CC Balaeopteridae; Balaeoptera.
 CX NCBI_TaxID=3768;
 RN [1]
 RP SEQUENCE.
 RA Ishihara Y., Saito T., Ito Y., Fujino M.
 RC "Structure of sperm- and sei-whale insulins and their breakdown by
 RC whale pepsin."
 RL Nature 181:1468-1479(1958).
 CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
 CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
 CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
 CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
 CC -!- SURUNIT: HETEROIDIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
 CC DISULFIDE BONDS.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
 CC P-R: A01582; INWHIS.
 DR HSSP: P01317; IAPH.
 DR InterPro: IPR004825; Ins/IGF/relax.
 DR PROSITE: PS00262; INSULIN; 1.
 KW Insulin family; Hormone; Glucose metabolism.
 FT CHAIN 1 30 INSULIN B CHAIN.
 FT NON CONS 30 31
 FT CHAIN 31 51 INSULIN A CHAIN.
 FT DISULFID 7 37 INTERCHAIN.
 FT DISULFID 19 50 INTERCHAIN.
 FT DISULFID 36 41

```

Boxidae; Caprinae; Capra.
NCBI_TaxID=3925;
[1]
SEQUENCE.
MEDLINE-66160119; PubMed-5949593;
Smith L.F.;
Species variation in the amino acid sequence of Insulin."
Am J Med. 40:662-666(1966).
FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND FATY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE CYCLE, AND GLYCOCEN SYNTHESIS IN LIVER.
SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO DISULFIDE BONDS.
CELLULAR LOCATION: Secreted.
SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
PIR: A01886; INCT.
HSSP: I07417; IAPH.
InterPro: IPRO04825; Ins/IGF/relax.
SMART: SMC0078; IGFP_1.
PROSITE: PS00282; INSGRN: 1.
Insulin Family; Hormone; Glucose metabolism.
CHAIN 1 36
INSULIN B CHAIN.
NON_CONS 30 31
CHAIN 31 51
INSULIN A CHAIN.
DISULFID 31 51
INTERCHAIN.
DISULFID 19 50
INTERCHAIN.
DISULFID 36 41
DISULFID 36 41
SEQUENCE 51 AA: 5692 MW: 5007B50CDM4E7DD CRC84;

Query Match      89.6%; Score 263.5; DB 1; Length 57;
Best Local Similarity 90.4%; pred. No. 8.2e-26;
Matches 47; Conservative 1; Mismatches 3; Indels 1; Gaps 0;

CY    1 FVNHCLGCSHLVEA-YLVCGRGFFTPKTRGVGECCCTSCSYQLENVCN 52
       ||||||||| | ||||||||| | |||..| | |||||
EB    1 FVNHCLGCSHALYLVCGRGFFTPKA-GIVECCAGVCSYQLENVCN 51

RESULT 11
INS_FIG
TC   INS_FIG STANDARD: PRT, 108 AA.
PC135; Q9TJ35;
21-JUL-1986 (Rel. 01, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
D8     Insulin precursor.
GN     INS.
US scrofa (pig);
EC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
MM Malia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
NCBI_TaxID=9823;
[1]
SEQUENCE FROM N.A.
Han X.G., Tuck B.E.;
KT "Complete porcine preproinsulin cDNA sequence.";
Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
[2]
SEQUENCE FROM N.A.
STRAIN-Large white;
XO MFELNP-2213598; PubMed=12140686;
FA Amarger V., Nguyen M., Laere A.S., Braunschweig M., Nezer C.,
EA Georges M., Andersson L.;
FA "Comparative sequence analysis of the INS-IGF2-HI9 gene cluster in pigs.";
ET Mamn, Genome 13:388-398(2002).
[3]
SEQUENCE OF 25-108
EX MEDLINP-68285485; PubMed=5657063;
FA Chance R.E., Ellis R.M., Bromer W.W.;
KT "Porcine proinsulin: characterization and aminic acid sequence.";
Science 161:165-167(1968).
[4]

```

RP REVISION TO 59.
 RA Chance R.E.
 RL Submitted (JUL-1970) to the PIR data bank.
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (1.3 ANGSTROMS).
 RA Blundell I.L., Dodson G.G., Hodgkin D., Moras D.
 RL "Insulin. The structure to the crystal and its reflection in
 chemistry and biology."
 RL Adv. Protein Chem. 26:279-402(1972).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
 RA Isaacs N.W., Adarwa R.C.
 RL "Experience with fast Fourier least squares in the refinement of the
 crystal structure of rhombohedral 2-zinc insulin at 1.5-A
 resolution."
 RL Acta Crystallogr. A 34:782-791(1978).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (1.3 ANGSTROMS).
 RA Baker E.N., Blundell I.L., Cutfield J.F., Dodson G.G.,
 RA Dodson G.G., Crowfoot Hodgkin D.M., Hubbard R.E., Isaacs N.W.,
 RA Reynolds C.D., Sakabe K., Sakabe N., Vijayan N.M.
 RL "The structure of 220 pig insulin crystals at 1.5-A resolution."
 RL Philos. Trans. R. Soc. Lond., B, Biol. Sci. 319:369-436(1978).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
 RA MEDLINE-92126280; PubMed-1772543.
 RA Batschmidt P., Hansen F.B., Dodson G., Korber F.
 RL "Structure of porcine insulin cocrystallized with elipatine 7."
 RL Acta Crystallogr. B 47:975-986(1991).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY.
 RA MEDLINE-91224250; PubMed-2925410.
 RA Badger G., Harris M.R., Reynolds C.D., Evans A.C., Dodson G.G.,
 RA Dodson G.G., North A.C.F.
 RL "Structure of the pig insulin dimer in the cubic crystal."
 RL Acta Crystallogr. B 47:127-136(1991).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (1.35 ANGSTROMS).
 RA Zhao J.-S., Wan Z.-L., Chang W.-K., Liang D.-Q.
 RL "Structure of monomeric porcine Des¹⁻³²-B2 aspartate peptide (326-546)
 insulin at 1.55-A resolution."
 RL Acta Crystallogr. B 53:507-512(1997).
 CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
 CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
 CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
 CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
 CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
 CC DISULFIDE BONDS.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
 CC -!- DATABASE: NAME-Protein Spotlight.
 CC NOTE=Issue 9 of April 2003.
 CC WWW="http://www.expasy.org/spotlight/articles/spotlight.html"
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.sib.ch/auction6/>
 CC or send an email to license@sib.ch).
 CC
 DR EMBL: AF064555; AAC77920.; ALT_INIT.
 DR EMBL: AY044828; AAL69550.;
 DR PDB: 3INS; 09-JAN-89.
 DR PDB: 4INS; 31-JUL-94.
 DR PDB: 6INS; 31-JAN-94.
 DR PDB: 7INS; 31-JAN-94.
 DR PDB: 9INS; 15-OCT-91.
 DR PDB: 112A; 15-OCT-91.
 DR PDB: 112B; 15-OCT-91.
 DR PDB: 2TC1; 29-JAN-96.

DR PDB: 1NEV; 29-JAN-96.
 DR PDB: 3MCH; 29-JAN-96.
 DR PDB: 1DEI; 16-JUN-97.
 DR PDB: 1SDB; 01-APR-98.
 DR PDB: 1WAV; 28-FEB-97.
 DR PDB: 1ZEE; 16-FEB-99.
 DR PDB: 12NI; 28-JAN-98.
 DR PDB: 12NJ; 28-JAN-98.
 DR InterPro: IPR004825; Ins/IGF/relax.
 DR Pfam: PF00049; Insulin; 1.
 DR SMART: SM00578; IIGF; 1.
 DR PROSITE: PS00262; INSULIN; 1.
 KW Insulin family; Hormone; Glucose metabolism; Signal; 3D-structure.
 FT SIGNAL; 24
 FT CHAIN; 25 54 INSULIN B CHAIN.
 FT PROPEP; 57 85 C-PEPTIDE.
 FT CHAIN; 86 106 INSULIN A CHAIN.
 FT DISULFID; 31 94 INTERCHAIN.
 FT DISULFID; 43 107 INTERCHAIN.
 FT DISULFID; 93 98
 FT HELIX; 26 46
 FT STRAND; 48 48
 FT HELIX; 89 94
 FT HELIX; 100 106
 FT STRAND; 107 107
 SC SEQUENCE 108 AA; 11671 MW; CB4491R429658EBE CRC64;
 Query Match 89.5%; Score 263; DB 1; Length 108;
 Blast Local Similarity 60.7%; Pred. No. 1.9e-25;
 Matches 51; Conservative 0; Mismatches 1; Indels 32; Gaps 1;
 QY 1 FVNHQHGSHLVEALYVCGEGRFFYPKTI----- 30
 DB 25 FVNHQHGSHLVEALYVCGEGRFFYPKARREAPNGAVELGGISGQAALAEQPP 84
 QY 31 ---KGIVEQCCTSGSLYLENYCN 52
 DB 25 QKRGIVEQCCTSGSLYLENYCN 108
 RESULT 12
 INS_RABIT STANDARD PRI: 110 AA.
 AC P01311;
 DE 21-JUL-1986 (rel. 01, Created)
 DE 21-FEB-1996 (rel. 33, Last sequence update)
 DE 21-OCT-1996 (rel. 34, Last annotation update)
 DE Insulin precursor.
 GN INS.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 ON NCBI_TaxID=9996;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=New Zealand white; TISSUE=Pancreas.
 RA MEDLINE-94179240; PubMed-8132571;
 RA Devaskar S.U., Giddings S.J., Rajakumar P.A., Carnaghi L.R.,
 RA Menon S.K., Zahm D.S.;
 RL "Insulin gene expression and insulin synthesis in mammalian neuronal
 cells."
 RL J. Biol. Chem. 269:8445-8454(1994).
 RN [2]
 RP SEQUENCE OF 25-54 AND 90-110.
 RX MEDLINE-66160119; PubMed-5949593;
 RA Smith L.F.;
 RL "Species variation in the amino acid sequence of insulin."
 RL Am. J. Med. 40:662-666(1966).
 RN [3]
 RP SEQUENCE OF 56-110 FROM N.A.
 RA Giddings S.J., Carnaghi L.R., Devaskar S.U.;
 RL Submitted (APR-1991) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT

```

CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -1- SUBUNIT: HETERO DIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch.)
CC
CC EMBL: U33610; AAA15033.1;
CC EMBL: M61253; AAA17540.1;
CC PIR: A53438; INPB.
CC HSSP: P01308; ITHM.
CC InterPro: IPR004825; Ins/IGF/relax.
CC SMART: SMC0078; IIGF.
CC PROSITE: PS00262; INSULIN.
CC Insulin family: Hormone; Glucose metabolism; Signal.
CC SIGNAL 1 24 INSULIN B CHAIN.
CC CHAIN 25 54 C PEPTIDE.
CC PROPEP 57 87 INSULIN A CHAIN.
CC CHAIN 90 110 INSULIN A CHAIN.
CC DISULFID 33 96 INTERCHAIN.
CC DISULFID 43 109 INTERCHAIN.
CC DISULFID 95 100 INTERCHAIN.
CC CONFLICT 83 85 E -> Y (IN REF. 3).
CC SSSEQUENCE 110 AA: 11818 MW: 826297585D77EAB CRC64:
CC
CC Query Match 89.5% Score 263; DB 1; Length 110;
CC Best Local Similarity 59.3%; Pred. No. 1.9e-25;
CC Matches 51; Conservative 1; Mismatches 0; Indels 34; Gaps 1;
CC
CC QY 1 FVNHLCGSHREVALYLVGGERGFYTPKTRREVERELVQGAELGGPGAGSLQPSALEL 84
CC 25 FVNHLCGSHREVALYLVGGERGFYTPKTRREVERELVQGAELGGPGAGSLQPSALEL 84
CC
CC DB 25 FVNHLCGSHREVALYLVGGERGFYTPKTRREVERELVQGAELGGPGAGSLQPSALEL 84
CC
CC QY 31 ----RGIVGCGCTSGSLYOLENYCN 52
CC 110 AA: 11818 MW: 826297585D77EAB CRC64:
CC DB 85 ALOKRGIVEGCTSGSLYOLENYCN 110
CC
CC RESULT 13
CC INS_SPETR STANDARD: PRT: 110 AA.
CC ID INS_SPETR
CC AC Q91X13;
CC DT 15-SEP-2003 (Rel. 42, Created)
CC DT 15-SEP-2003 (Rel. 42, Last sequence update)
CC DT 15-SEP-2003 (Rel. 42, Last annotation update)
CC DE Insulin precursor.
CC INS.
CC OS Spermophilus tridecemlineatus (Thirteen-lined ground squirrel).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Scuridae; Scuriinae;
CC Spermophilus.
CC NCBI_TaxID=43179;
CC RN 11;
CC RP SEQUENCE FROM N.A.
CC RC Tissue-Pancreas;
CC RA Fredrea M.M., Buck M.J., Ghanayori J., Squire P.D., Andrews K.L.;
CC RT "Regulation of PDK4 expression in a hibernating mammal";
CC RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -1- SUBUNIT: HETERO DIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO

```

```

CC DISULFIDE BONDS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch.)
CC
CC EMBL: AY338604; AAK72558.1;
CC HSSP: P01308; ILNP.
CC InterPro: IPR004825; Ins/IGF/relax.
CC PIR: P00049; INSULIN.
CC SMART: PR00277; INSULINB.
CC PROSITE: PS00262; INSULIN.
CC Insulin family: Hormone; Glucose metabolism; Signal.
CC SIGNAL 24 BY SIMILARITY.
CC CHAIN 25 54 INSULIN B CHAIN.
CC PROPEP 57 87 C PEPTIDE.
CC CHAIN 90 110 INSULIN A CHAIN.
CC DISULFID 31 94 INTERCHAIN (BY SIMILARITY).
CC DISULFID 43 109 INTERCHAIN (BY SIMILARITY).
CC DISULFID 95 100 BY SIMILARITY.
CC SSSEQUENCE 110 AA: 12004 MW: 4511768D6522BEE5 CRC64:
CC
CC Query Match 89.5% Score 263; DB 1; Length 110;
CC Best Local Similarity 59.3%; Pred. No. 1.9e-25;
CC Matches 51; Conservative 1; Mismatches 0; Indels 34; Gaps 1;
CC
CC QY 1 FVNHLCGSHREVALYLVGGERGFYTPKTRREVERELVQGAELGGPGAGSLQPSALEL 84
CC 25 FVNHLCGSHREVALYLVGGERGFYTPKTRREVERELVQGAELGGPGAGSLQPSALEL 84
CC
CC DB 25 FVNHLCGSHREVALYLVGGERGFYTPKTRREVERELVQGAELGGPGAGSLQPSALEL 84
CC
CC QY 31 ----RGIVGCGCTSGSLYOLENYCN 52
CC 110 AA: 11818 MW: 826297585D77EAB CRC64:
CC DB 85 ALOKRGIVEGCTSGSLYOLENYCN 110
CC
CC RESULT 14
CC INS_FELCA STANDARD: PRT: 51 AA.
CC ID INS_FELCA
CC AC P06306;
CC DT 01-JAN-1988 (Rel. 06, Created)
CC DT 01-JAN-1988 (Rel. 06, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Insulin.
CC INS.
CC OS Felis silvestris catus (Cat).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
CC NCBI_TaxID=9685;
CC RN 11;
CC RP SEQUENCE.
CC RA MEDLINE=86214076; PubMed=3518635;
CC RA Hallden G., Gavell G., Matt V., Jorcnvall H.;
CC RT "Characterization of cat insulin.";
CC ML Arch. Biochem. Biophys. 247:20-27(1986).
CC -1- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -1- SUBUNIT: HETERO DIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC PIR: A01588; INCT.
CC HSSP: P01317; IAPH.
CC InterPro: IPR004825; Ins/IGF/relax.
CC PRINTS: PR00277; INSULINB.

```


GenCore version 5.1.4
Copyright (c) 1993 2003 Computer, Ltd.

OM protein - protein search, using sw Model:

Run on: September 15, 2003, 11:54:30 : Search time 27.314 seconds
(with 2003 alignments)
660,520 Million Cells updated/sec

Title: US-09-423-100-5

Perfect score: 274

Sequence: 1 FVNHQCGSHLV-ALYVCG.....VEQCCTSCSYOLENYCN 52

Scoring table: ELOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_humani:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mice:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_p-anti:*
- 11: sp_protist:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_yeast:*
- 16: sp_bacteriophage:*
- 17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	267	90.8	110	Q8HXV2	Q8HXV2 pongo pygmaeus
2	265	89.5	110	Q91X13	Q91X13 pongo pygmaeus
3	251	85.4	110	Q8WKK6	Q8WKK6 pongo pygmaeus
4	219.5	74.7	106	Q918Q7	Q918Q7 pongo pygmaeus
5	201.5	68.5	110	Q98TA6	Q98TA6 pongo pygmaeus
6	195.5	66.5	108	Q90DP5	Q90DP5 brachycephalus
7	195.5	66.5	108	Q902N4	Q902N4 cala cala
8	195	66.3	113	Q98TB0	Q98TB0 chitara cha
9	193.5	65.8	110	Q98ZY1	Q98ZY1 bicodon alios
10	191.5	65.1	111	Q98TA7	Q98TA7 osteocephalus
11	187.5	63.8	87	Q98TA9	Q98TA9 acanthomys
12	186.5	63.4	108	Q98TB1	Q98TB1 catostomus
13	185.5	63.1	91	Q98TB2	Q98TB2 ameleptites
14	146	49.7	65	Q8H281	Q8H281 acilia gor
15	146	49.7	65	Q8H280	Q8H280 pongo pygmaeus
16	144	49.0	207	Q90XD0	Q90XD0 cyprinid ca

17	144	49.0	215	13	Q73721	Q73721 tilapia sp.
18	143.5	48.8	132	13	Q8AV14	Q8AV14 petromyzon
19	141.5	48.1	163	13	Q93607	Q93607 paralichthy
20	141.5	48.1	162	13	Q73720	Q73720 oreochromis
21	141.5	48.1	182	13	Q42289	Q42289 oreochromis
22	141.5	48.1	182	13	Q79824	Q79824 oreochromis
23	141.5	48.1	185	13	Q57436	Q57436 paralichthy
24	141.5	48.1	186	13	Q93527	Q93527 paralichthy
25	141	48.0	185	13	Q9Y157	Q9Y157 acanthopagr
26	140.5	47.8	116	13	Q91181	Q91181 oncorhynch
27	140.5	47.8	117	13	Q91476	Q91476 salmo salar
28	140.5	47.8	145	13	Q91475	Q91475 salmo salar
29	140.5	47.8	149	13	Q91231	Q91231 oncorhynch
30	140.5	47.8	155	13	Q91162	Q91162 oncorhynch
31	140.5	47.8	161	13	Q91230	Q91230 oncorhynch
32	140.5	47.8	184	13	Q42336	Q42336 myoxocephal
33	140.5	47.8	188	13	Q81288	Q81288 oncorhynch
34	140.5	47.8	188	13	Q91965	Q91965 oncorhynch
35	140	47.6	210	13	Q91443	Q91443 squallus aca
36	139	47.3	212	13	Q8J184	Q8J184 brachydanio
37	139	47.3	215	13	Q42429	Q42429 lates calca
38	138.5	47.1	123	6	Q8MJ15	Q8MJ15 sus scrofa
39	138.5	47.1	153	13	Q93380	Q93380 melegritis g
40	137.5	46.8	62	13	Q91AA0	Q91AA0 carassius a
41	137.5	46.8	117	13	Q91914	Q91914 ctenopharyn
42	137.5	46.8	149	6	Q9MYX4	Q9MYX4 bos indicus
43	137.5	46.8	161	13	Q90VV9	Q90VV9 brachydanio
44	137.5	46.8	161	13	Q9PWK2	Q9PWK2 carassius a
45	137.5	46.8	161	13	Q96SR6	Q96SR6 megalobrama

ALIGNMENTS

RESULT 1:
Q8HXV2 PRELIMINARY: PRT: 110 AA.
ID Q8HXV2
AC Q8HXV2
DC Q1-MAR-2003 (TREMBLrel. 23, Created)
DT Q1-MAR-2003 (TREMBLrel. 23, Last sequence update)
DE Q1-MAR-2003 (TREMBLrel. 23, Last annotation update)
DF Insulin precursor.
GR INS.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI TaxID:9600;
ER 1

RP SEQUENCE FROM N.A.
RA Stead J.C.H., Jeffreys A.J.:
RT "Haplotype diversity at the insulin region."
RL Submitted (JUL-2002) to the EMBL/GenBank/CDDB databases.
PR EMBL: AV137503; AAC06537; ...
SQ SEQUENCE 110 AA, 12018 MW: 22023209452048 CRC64;

Query Match 90.8%; Score 267; DB 6; Length 110;
Best local Similarity 62.5%; Prod. No. 1e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 2;

QY 1 FVNHQCGSHLV-ALYVCG.....VEQCCTSCSYOLENYCN 52
|||||
DB 25 FVNHQCGSHLV-ALYVCG.....VEQCCTSCSYOLENYCN 110

QY 31 ----SGIVEQCCTSCSYOLENYCN 52
|||||
DB 85 SLOKRGVEQCCTSCSYOLENYCN 110

RESULT 2
Q91X13 PRELIMINARY: PRT: 110 AA.
ID Q91X13
AC Q91X13
DT Q1-DEC-2001 (TREMBLrel. 19, Created)

```

QY 31 ----RGIVEQCCISGLYOLENYCN 52
      ||| ||| ||| ||| ||| ||| ||| |||
DB 85 FLQRGIVEQCCASVCSLYOLEHYN 110

RESULT 4
Q91507 PRELIMINARY: PRT; 106 AA.
AC Q91507:
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Preproinsulin.
OS Kana pipiens (Northern leopard frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Anura; Neobatrachia; Ranidae; Rana;
OC NCBI_TaxID=8404;
LN 11
KF SEQUENCE FROM N.A.
RX MEDLINE=20362507; PubMed=10818274;
RX Irwin D.M., Sivaraiah P.;
KT "Proinsulin cDNAs from the leopard frog, Rana pipiens: evolution of
KT proinsulin processing".
KL Comp. Biochem. Physiol. 125B:405-410 (2000).
CC -! SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -! SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AF227167; AAF87205.1; -.
DR HSSP: P01315; 1SDR.
DA InterPro: IP004625; Ins/IGF/relax.
DA Pfam: PF00044; Insulin_1.
DA PRINTS: PR02377; INSULIN5.
DA SMART: SM00278; IIGF_1.
DA PROSITE: PS00262; INSULIN_1.
SQ SEQUENCE 106 AA; 12183 MW; 3AB7CEC70217F92 CRC64;

Query Match 74.7% Score 219.5; D9 I3; Length 106;
Best local similarity 49.4%; Pred. No. 3e-22;
Matches 41; Conservative 7; Mismatches 4; Indels 31; Gaps 0

QY 1 FVNHGLCSHLYEALYNCGEFTPTKTR-----
      ||| ||| ||| ||| ||| ||| ||| |||
QY 24 FDNQYICGSHLYEALVWVGDRGFYPSRRDLQELVNGLQSSLEDEMVOVSQAF
      ||| ||| ||| ||| ||| ||| ||| |||
QY 32 --GIVEQCCISGLYOLENYCN 52
      ||| ||| ||| ||| ||| ||| |||
DB 84 KPGIVEQCCHNTCSLYDLENYCN 106

RESULT 5
Q98TAE PRELIMINARY: PRT; 110 AA.
AC Q98TAE:
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Preproinsulin.
OS Pantodon buchholzi (Butterflyfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Pantodontidae; Pantodon.
OC NCBI_TaxID=8276;
LN 11
KF SEQUENCE FROM N.A.
RX MEDLINE=21033577; PubMed=11306171;
RX A. Mahrooki A.A., Irwin D.M., Graham L.C., Youson J.H.;
KT "Molecular cloning of preproinsulin cDNAs from several
KT osteoglossomorphs and a cyprinid.".
DA Mol. Cell. Endocrinol. 174:51-58 (2001).
CC -! SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -! SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AF199586; AAK28712.1; -.
DR HSSP: P01308; 1HIS.

```



```

DR InterPro: IPR004625; Ins/IGF/relax.
DR Pfam: PF00049; Insulin 1.
DR SMART: SM00378; IIGF; 1.
DR PROSITE: PS00262; INSULIN 1.
SQ SEQUENCE 110 AA; 12424 MW; HDECCDQSGSSE72P06 GR64;

Query Match 68.5%; Score 201.5; DP 13; Length 110.
Best Local Similarity 43.1%; Pred. No. 5.8e-19;
Matches 37; Conservative 8; Mismatches 5; Indels 3; Gaps 11

QY 3 NOHCGSHLVKALYLVCGRGGFFYTPK-----T 30
RA Hhatacharya S., Roy S.S., Dasgupta S., Ravikumar L., Mukherjee M.,
BA Bandyopadhyaya S., Wakabayashi K.,
RT "A new cell secreting insulin."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AF373021; AAK51568.1; -.
DR HSP: P01308; IIGF; 1.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin; 1.
DR SMART: SM00277; INSULIN.
DR PROSITE: PS00262; IIGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
SQ SEQUENCE 108 AA; 11881 MW; D713026E22EF5D59 GR64;

Query Match 66.5%; Score 195.5; DP 13; Length 108;
Best Local Similarity 45.1%; Pred. No. 5.8e-19;
Matches 37; Conservative 5; Mismatches 7; Indels 3; Gaps 1;

QY 4 QHICGSHLVKALYLVCGRGGFFYTPK-----T 30
DB 27 QHICGSHLVKALYLVCGRGGFFYTPKRDVDPLMGFLPKSAQETEVADFADKHAIVRX 86

QY 31 RGIVEQCCTSLCSLYOLENYCN 52
DB 87 RGIVEQCCHKPCSPFELONYCN 108

RESULT 8
Q902R0 Q902R0 PRELIMINARY; PRI: 111 AA.
AC Q98TB0;
DT 01-JUN-2001 (TRENDELrel. 17, Created)
DT 01-JUN-2001 (TRENDELrel. 17, Last sequence update)
DT 01-MAR-2003 (TRENDELrel. 23, Last annotation update)
DB Preproinsulin (Fragment);
OS Capitula chitila (clown knifefish);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Notopteridae; Chitila.
OX NCBI_TaxID=112263;
RN 111
SF SEQUENCE FROM N.A.
ME MEDLINE: 2263577; PubMed: 11306171;
RA A.; Mahrouki A.A.; Irwin D.M.; Graham L.C.; Youson J.H.;
RT "Molecular cloning of preproinsulin cDNAs from several
OS osteoglossomorphs and a cyprinid."
CC -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AF19586; AAK28710.1; -.
DR HSP: P01308; IIGF; 1.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin; 1.
DR SMART: SM00078; IIGF; 1.
DR NCNTER 111
SQ SEQUENCE 111 AA; 12483 MW; 247CA443376329F GR64;

Query Match 66.3%; Score 195; DP 13; Length 111;
Best Local Similarity 44.2%; Pred. No. 7e-19;
Matches 38; Conservative 3; Mismatches 9; Indels 36; Gaps 1;

QY02N4 Q902N4 PRELIMINARY; PRI: 108 AA.
AC Q902N4;

```

```

DT 01-DEC-2001 (TRENDELrel. 19, Created)
DT 01-DEC-2001 (TRENDELrel. 19, Last sequence update)
DT 01-MAR-2003 (TRENDELrel. 23, Last annotation update)
DB Preproinsulin.
OS Capitula chitila (clown knifefish);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteiiformes; Cypriniformes;
OC Cyprinidae; Catfish.
OX NCBI_TaxID=72446;
RN 111
SF SEQUENCE FROM N.A.
ME MEDLINE: 2263577; PubMed: 11306171;
RA Hhatacharya S., Roy S.S., Dasgupta S., Ravikumar L., Mukherjee M.,
BA Bandyopadhyaya S., Wakabayashi K.,
RT "A new cell secreting insulin."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AF373021; AAK51568.1; -.
DR HSP: P01308; IIGF; 1.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin; 1.
DR SMART: SM00277; INSULIN.
DR PROSITE: PS00262; IIGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
SQ SEQUENCE 108 AA; 11881 MW; D713026E22EF5D59 GR64;

Query Match 66.5%; Score 195.5; DP 13; Length 108;
Best Local Similarity 45.1%; Pred. No. 5.8e-19;
Matches 37; Conservative 5; Mismatches 7; Indels 3; Gaps 1;

QY 4 QHICGSHLVKALYLVCGRGGFFYTPK-----T 30
DB 27 QHICGSHLVKALYLVCGRGGFFYTPKRDVDPLMGFLPKSAQETEVADFADKHAIVRX 86

QY 31 RGIVEQCCTSLCSLYOLENYCN 52
DB 87 RGIVEQCCHKPCSPFELONYCN 108

RESULT 8
Q902R0 Q902R0 PRELIMINARY; PRI: 111 AA.
AC Q98TB0;
DT 01-JUN-2001 (TRENDELrel. 17, Created)
DT 01-JUN-2001 (TRENDELrel. 17, Last sequence update)
DT 01-MAR-2003 (TRENDELrel. 23, Last annotation update)
DB Preproinsulin (Fragment);
OS Capitula chitila (clown knifefish);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Notopteridae; Chitila.
OX NCBI_TaxID=112263;
RN 111
SF SEQUENCE FROM N.A.
ME MEDLINE: 2263577; PubMed: 11306171;
RA A.; Mahrouki A.A.; Irwin D.M.; Graham L.C.; Youson J.H.;
RT "Molecular cloning of preproinsulin cDNAs from several
OS osteoglossomorphs and a cyprinid."
CC -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AF19586; AAK28710.1; -.
DR HSP: P01308; IIGF; 1.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin; 1.
DR SMART: SM00078; IIGF; 1.
DR NCNTER 111
SQ SEQUENCE 111 AA; 12483 MW; 247CA443376329F GR64;

Query Match 66.3%; Score 195; DP 13; Length 111;
Best Local Similarity 44.2%; Pred. No. 7e-19;
Matches 38; Conservative 3; Mismatches 9; Indels 36; Gaps 1;

QY02N4 Q902N4 PRELIMINARY; PRI: 108 AA.
AC Q902N4;

```

```

QY 3 NQHLGSHLVEALYLVCGERGFFYTPKT----- 30
      ||||| ||| ||||| ||||| ||
DB 26 NQHLGSHLVEALYLVCGERGFFYTPKMDKRAEPLGLGFLSPKSGLENEVDYKPEQAEVK 86
      ||||| ||| ||||| ||||| ||
QY 30 ---TRGIVEQCCTSIQSLYLENYCN 52
      ||||| ||| ||||| ||
DB 26 VKMKRGIVEQCCHRPNCNFDLQNYCN 112
      ||||| ||| ||||| ||

RESULT 9
Q902Y1 ID Q902Y1 PRELIMINARY: PRT: 111 AA.
AC Q902Y1:
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE Preproinsulin (Fragment)
OS Hiodon alosoides (holdeye).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Hiodontidae; Hiodon.
OX NCBI_TaxID=54924.
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=21203577; PubMed=11306171.
RA Al-Mahrouki A.A., Irwin D.M., Graham L.C., Youson J.H.;
RT "Molecular cloning of preproinsulin cDNAs from several
osteoglossomorphs and a cyprinid."
RL Mol. Cell. Endocrinol. 174:51-58(2001).
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AF282496; AAK24064.1; -.
DR HSSP: PG1308; I1NF.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin; 1.
DR PRINTS: PG0277; INSULIN.
DR SMART: SMC0076; IIGF; 1.
DR PROSITE: PS00252; INSULIN; 1.
FT NON_TER 112 112
FT NON_TER 112 112
SQ SEQUENCE 110 AA: 12345 MW: 85552.52 DSE CRC64:
      MEDLINE=21203577; PubMed=11306171.
      Query Match 65.8%; Score 193.5; DB 13; Length 110;
      Best Local Similarity 42.8%; Pred. No. 1,1e-18;
      Matches 36; Conservative 7; Mismatches 7; Indels 35; Gaps 1;

QY 3 NQHLGSHLVEALYLVCGERGFFYTPKT----- 30
      ||||| ||| ||||| ||||| ||
DB 26 NQHLGSHLVEALYLVCGERGFFYTPKMDKRAEPLGLGFLSPKSGLENEVDYKPEQAEVK 86
      ||||| ||| ||||| ||||| ||
QY 31 ---RGIVEQCCTSIQSLYLENYCN 52
      ||||| ||| ||||| ||
DB 86 VKMKRGIVEQCCHRPNCNFDLQNYCN 110
      ||||| ||| ||||| ||

RESULT 10
Q987A7 ID Q987A7 PRELIMINARY: PRT: 111 AA.
AC Q987A7:
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Preproinsulin (Fragment)
OS Osteoglossum bicirrhosum (silver arowana).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Osteoglossidae; Osteoglossum.
OX NCBI_TaxID=109271.
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=21203577; PubMed=11306171.
RA Al-Mahrouki A.A., Irwin D.M., Graham L.C., Youson J.H.;
RT "Molecular cloning of preproinsulin cDNAs from several
osteoglossomorphs and a cyprinid."

```

```

RD Mol. Cell. Endocrinol. 174:51-58(2001).
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AF199589; AAK28713.1; -.
DR HSSP: F01315; I1PJ.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin; 1.
DR PRINTS: PG0277; INSULIN.
DR SMART: SMC0076; IIGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
FT NON_TER 111 111
FT NON_TER 111 111
SQ SEQUENCE 111 AA: 12491 MW: AC9E19D2D4866D2C CRC64:
      Query Match 65.1%; Score 191.5; DB 13; Length 111;
      Best Local Similarity 41.2%; Pred. No. 2,1e-18;
      Matches 35; Conservative 10; Mismatches 5; Indels 35; Gaps 1;

QY 3 NQHLGSHLVEALYLVCGERGFFYTPKT----- 30
      ||||| ||| ||||| ||||| ||
DB 27 NQHLGSHLVEALYLVCGERGFFYTPKMDKRAEPLGLGFLSPKSGLENEVDYKPEQAEVK 86
      ||||| ||| ||||| ||||| ||
QY 31 ---RGIVEQCCTSIQSLYLENYCN 52
      ||||| ||| ||||| ||
DB 87 VKMKRGIVEQCCHRPNCNFDLQNYCN 111
      ||||| ||| ||||| ||

RESULT 11
Q987A7 ID Q987A7 PRELIMINARY: PRT: 87 AA.
AC Q987A7:
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE Preproinsulin (Fragment)
OS Gnathonemus petersi.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Mormyridae; Gnathonemus.
OX NCBI_TaxID=42645.
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=21203577; PubMed=11306171.
RA Al-Mahrouki A.A., Irwin D.M., Graham L.C., Youson J.H.;
RT "Molecular cloning of preproinsulin cDNAs from several
osteoglossomorphs and a cyprinid."
RL Mol. Cell. Endocrinol. 174:51-58(2001).
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AF199587; AAK28711.1; -.
DR HSSP: F01308; I1IS.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin; 1.
DR SMART: SMC0078; IIGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
FT NON_TER 87 87
FT NON_TER 87 87
SQ SEQUENCE 87 AA: 9874 MW: FF448ED35D2453F5 CRC64:
      Query Match 63.8%; Score 187.5; DB 13; Length 87;
      Best Local Similarity 42.9%; Pred. No. 5,8e-18;
      Matches 36; Conservative 5; Mismatches 8; Indels 35; Gaps 1;

QY 4 QHLGSHLVEALYLVCGERGFFYTPKT----- 30
      ||||| ||| ||||| ||||| ||
DB 4 QHLGSHLVEALYLVCGERGFFYTPKMDKRAEPLGLGFLSPKSGLENEVDYKPEQAEVK 63
      ||||| ||| ||||| ||||| ||
QY 31 ---RGIVEQCCTSIQSLYLENYCN 52
      ||||| ||| ||||| ||
DB 64 VKMKRGIVEQCCHRPNCNFDLQNYCN 87
      ||||| ||| ||||| ||

RESULT 12
Q967A7 ID Q967A7 PRELIMINARY: PRT: 111 AA.
AC Q967A7:
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Preproinsulin (Fragment)
OS Osteoglossum bicirrhosum (silver arowana).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Osteoglossidae; Osteoglossum.
OX NCBI_TaxID=109271.
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=21203577; PubMed=11306171.
RA Al-Mahrouki A.A., Irwin D.M., Graham L.C., Youson J.H.;
RT "Molecular cloning of preproinsulin cDNAs from several
osteoglossomorphs and a cyprinid."

```

```

ID Q98TB2: PRELIMINARY: PRT: 128 AA.
AC Q98TB2:
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Preproinsulin (Fragment).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
OC Catostomidae; Catostomus.
OX NCBI_TaxID=9971;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21203577; PubMed-11306171;
RA Al-Mahrouki A.A., Irwin D.M., Graham D.C., Yanson J.H.;
RT "Molecular cloning of preproinsulin cDNAs from several
RT osteoglossomorphs and a cyprinid."
RL Mol. Cell. Endocrinol. 174(S1-58(2001)).
CC -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AF199584; AAK28708.1; -.
DR HSSP: P01308; ILPH.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF000049; Insulin; 1.
DR SMART: SM00078; IIGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
FT NON_TER 108
FT SEQUENCE 108 AA; 11873 MW; E426310696FBAF06 CRC64;

Query Match 63.4%; Score 186.5; DB 13; Length 108;
Best Local Similarity 43.9%; Pred. No. 9.6e-18;
Matches 36; Conservative 4; Mismatches 9; Indels 3; Gaps 1;

QY 4 QHLGGSHLVFALYLVCGRGFFYTPK-----T 40
DE 27 QHLGGSHLVFALYLVCGRGFFYTPKTRRRAEDLVGQVGVGGAGSLOPLALEGSLQKRG 60

QY 31 RGIVGQCTSLCSLYOLENYCN 52
DE 27 RGIVGQCTSLCSLYOLENYCN 52

QY 37 RGIVGQCTSLCSLYOLENYCN 52
DE 27 RGIVGQCTSLCSLYOLENYCN 52

RESULT 13
Q98TB2: PRELIMINARY: PRT: 91 AA.
AC Q98TB2:
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Preproinsulin (Fragment).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
OC Catostomidae; Ambloplites.
OX NCBI_TaxID=103273;
RN [1]
RP SEQUENCE FROM N.A.
RX Al-Mahrouki A.A., Irwin D.M., Yanson J.H.;
RT "Molecular cloning of preproinsulin cDNA from the rock bass."
RL Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
CC -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AF199584; AAK28708.1; -.
DR HSSP: P01308; ILPH.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF000049; Insulin; 1.
DR SMART: SM00078; IIGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
FT NON_TER 91
FT SEQUENCE 91 AA; 10100 MW; E36C882560C65D39 CRC64;

```

```

Query Match 64.1%; Score 185.5; DB 13; Length 91;
Best Local Similarity 40.9%; Pred. No. 1.1e-17;
Matches 46; Conservative 5; Mismatches 8; Indels 39; Gaps 1;

QY 4 QHLGGSHLVFALYLVCGRGFFYTPK----- 29
DE 4 QHLGGSHLVFALYLVCGRGFFYTPKTRRRAEDLVGQVGVGGAGSLOPLALEGSLQKRG 63
QY 46 -----IRGIVGQCTSLCSLYOLENYCN 52
DE 46 -----IRGIVGQCTSLCSLYOLENYCN 52
QY 64 MEMVKKRGIVGQCTSLCSLYOLENYCN 91
DE 64 MEMVKKRGIVGQCTSLCSLYOLENYCN 91

RESULT 14
Q98TB2: PRELIMINARY: PRT: 65 AA.
AC Q98TB2:
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Insulin (Fragment).
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RX O'Duigain C., Tichy H., Klein J.;
RT "Molecular evolution in higher primates: gene specific and organism
RT specific characteristics."
RT Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL: AY092023; AAM76640.1; -.
FT NON_TER 1
FT NON_TER 65
FT SEQUENCE 65 AA; 6920 MW; B772017FD8BCABEA CRC64;

Query Match 49.7%; Score 146; DB 6; Length 65;
Best Local Similarity 47.7%; Pred. No. 2e-12;
Matches 31; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 7 QGSHLVFALYLVCGRGFFYTPK-----RG 32
DE 7 QGSHLVFALYLVCGRGFFYTPKTRRRAEDLVGQVGVGGAGSLOPLALEGSLQKRG 60

QY 43 IVEQC 37
DE 43 IVEQC 65

RESULT 15
Q98TB2: PRELIMINARY: PRT: 65 AA.
AC Q98TB2:
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Insulin (Fragment).
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX O'Duigain C., Tichy H., Klein J.;
RT "Molecular evolution in higher primates: gene specific and organism
RT specific characteristics."
RT Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL: AY092024; AAM76641.1; -.
FT NON_TER 1
FT NON_TER 65
FT SEQUENCE 65 AA; 6920 MW; B772017FD8BCABEA CRC64;

Query Match 49.7%; Score 146; DB 6; Length 65;

```

Best Local Similarity 47.7% Prod. No. 2c.12:
 Matches 31: Conservative 0; Mismatches 0; Indels 34; Gaps 1;
 QY 7 CGSHLVPAALVVGSGRGFFVTPKT-----EQL
 Db 1 CGSHLVPAALVVGSGRGFFVTPKT-----EQL
 QY 33 IVFQC 37
 Db 61 IVFQC 65

Search completed: September 15, 2003, 12:03:30
 Job time : 21.3154 secs

PT particularly for the production of human insulin -
XX
XX
XX Claim 13; Page 30; 46pp; English.
XX
CC This sequence represents a chimeric protein, hGH-mini-proinsulin.
CC This chimeric protein contains an N-terminal fragment of human growth
CC hormone (hGH) of the sequence given in AAY42855, a cleavable peptide
CC linker (AAY42857), and a human insulin precursor comprising insulin
CC A and B chains (AAY42859). The hGH portion of the chimeric protein acts
CC as an intramolecular chaperone (IMC) for the insulin precursor.
CC C-terminal Arg residue which enables the hGH portion of the
CC chimeric protein to be removed after folding has taken place. Production
CC of recombinant human insulin via an hGH-proinsulin chimeric protein can
CC provide human insulin with correctly linked cysteine bridges with
CC fewer necessary procedural steps, and hence resulting in a higher yield
CC of human insulin. The IMC sequences not only protect insulin sequences
CC from intracellular degradation by a microorganism host, but also promote
CC the folding of the fused insulin precursor, facilitate the solubility of
CC the fusion protein and decrease the intermolecular interactions among
CC the fusion proteins, thus allowing folding of the fused insulin precursor
CC at commercially useful high concentrations. The procedural steps of
CC cyanogen bromide cleavage, oxidative sulphytolysis and related
CC purification steps can thus be eliminated, along with the use of high
CC concentrations of mercaptan or the use of hydrophobic absorbent resins.
XX
XX Sequence 107 AA:
SQ
Query Match 100.0%; Score 587; DA 20; Length 107;
Best Local Similarity 100.0%; Pred. No. 49-49;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MEPTIPLSRLFDNAMIKAHRLHQLAFTDYQEEFAVIEKEOKYSFLQNP----- 49
DB 1 MEPTIPLSRLFDNAMIKAHRLHQLAFTDYQEEFAVIEKEOKYSFLQNPQTSLSFSESIP 60
OY 61 LQSHLYEALVCGGSGFFYTKRIGVEQCTSCSLYQLENYCN 107
DB 61 LQSHLYEALVCGGSGFFYTKRIGVEQCTSCSLYQLENYCN 107
RESULT 2
RAY42861
ID AAY42861 standard; protein: 150 AA.
AC AAY42861;
XX
XX (9-JAN-2000 (first entry)
XX Chimeric protein, SEQ ID 7.
XX Insulin; precursor; growth hormone; chaperone; intramolecular;
XX folding; conformation; chimeric protein; cleavable; recombinant;
XX production; yield.
XX Synthetic.
XX Homo sapiens.
XX WO9950302-A1.
XX 07-OCT-1999.
XX
XX 31-MAR-1998; 98WO-CN00052.
XX 31-MAR-1998; 98WO-CN00052.
XX (TONG-) TONGHUA GANTECH BIOTECHNOLOGY LTD.
XX Gan Z;
XX WPI; 1999-610839/52.
XX
XX New chimeric proteins containing human growth hormone fragment, used
PT particularly for the production of human insulin -
XX
XX Claim 14; Page 30-31; 46pp; English.
XX
CC This sequence represents a chimeric protein, which contains an
CC N-terminal fragment of human growth hormone (hGH) of the sequence given
CC in AAY42856, a cleavable peptide linker (AAY42857), and a human insulin
CC precursor comprising insulin A and B chains (AAY42859). The hGH portion
CC of the chimeric protein acts as an intramolecular chaperone (IMC) for
CC the insulin precursor, enabling it to fold correctly. The cleavable
CC peptide linker has a C-terminal Arg residue which enables the hGH portion
CC of the chimeric protein to be removed after folding has taken place.
CC Production of recombinant human insulin via an hGH-proinsulin chimeric
CC protein can provide human insulin with correctly linked cysteine bridges
CC with fewer necessary procedural steps, and hence resulting in a higher
CC yield of human insulin. The IMC sequences not only protect insulin
CC sequences from intracellular degradation by a microorganism host, but
CC also promote the folding of the fused insulin precursor, facilitate the
CC solubility of the fusion protein and decrease the intermolecular
CC interactions among the fusion proteins, thus allowing folding of the
CC fused insulin precursor at commercially useful high concentrations. The
CC procedural steps of cyanogen bromide cleavage, oxidative sulphytolysis
CC and related purification steps can thus be eliminated, along with the use
CC of high concentrations of mercaptan or the use of hydrophobic absorbent
CC resins.
XX
XX Sequence 150 AA:
SQ
Query Match 94.6%; Score 555.5; DA 20; Length 150;
Best Local Similarity 71.3%; Pred. No. 2.4e-39;
Matches 107; Conservative 0; Mismatches 0; Indels 43; Gaps 1;
OY 1 MEPTIPLSRLFDNAMIKAHRLHQLAFTDYQEEFAVIEKEOKYSFLQNP----- 49
DB 1 MEPTIPLSRLFDNAMIKAHRLHQLAFTDYQEEFAVIEKEOKYSFLQNPQTSLSFSESIP 60
OY 50 -----LGTGPRFVNOHLOGSHLYEALVCGGSR 77
DB 61 TISNREEQQKSNELLRTSIILOSLEPVCLGTGPRFVNOHLOGSHLYEALVCGGSR 120
OY 78 GFFYTPKTRIGVEQCTSCSLYQLENYCN 107
DB 121 GFFYTPKTRIGVEQCTSCSLYQLENYCN 150
RESULT 3
AAR96897
ID AAR96897 standard; protein: 115 AA.
XX AAR96897;
XX
XX 03-FEB-1997 (first entry)
XX SOD-proinsulin hybrid polypeptide.
XX Insulin; proinsulin; hybrid polypeptide; protein folding;
XX enzymatic cleavage; cyanogen bromide; sulphytolysis.
XX Homo sapiens.
XX WO9620724-A1.
XX 11-JUL-1996.
XX 29-DEC-1994; 94WO-US13268.
XX 29-DEC-1994; 94WO-US13268.
XX 10-JAN-1995; 95ZA-0000142.
XX (BIOT-) BIO-TECHNOLOGY GENERAL CORP.
XX Gcrecki M, Hartman JR, Mendelovitz S;
XX

CC host cells with pBAST-R results in the efficient expression of the
 CC proinsulin hybrid polypeptide, useful for human insulin production.
 CC The method produces recombinant human insulin identical to the
 CC natural hormone. Hazardous and cumbersome procedures involving
 CC cyanogen bromide and sulphitolysis to protect SH groups are avoided
 CC since the entire hybrid polypeptide folds efficiently to the native
 CC structure even with the leader attached and Cys unprotected.

XX Sequence 117 AA;

Query Match 51.3%; Score 304; DB 17; Length 117;
 Best Local Similarity 82.6%; Pred. No. 2,3e-18;
 Matches 57; Conservative 3; Mismatches 5; Indels 4; Gaps 2;

QY 43 YSFIONPIGT--GPRFVNHLGSHLVEALYLVCGERGFFYPTKTRGIVLQDLSLSS 56
 DB 45 HEFGNTEAGSTAGPRFVNHLGSHLVEALYLVCGERGFFYPTKTRGIVLQDLSLSS 108

QY 9% DYOLENYCN 107
 DB 10% DYOLENYCN 117

RESULT 6
 AAR71692
 ID AAR71692 standard; Protein: 117 AA.
 AC AAR71692;
 XX 25-MAR-2003 (updated)
 DT 20-NOV-1995 (first entry)
 XX
 XX Mating factor alpha 1-Insulin precursor Arg831.
 XX
 XX Human insulin precursor Arg831; diabetes; zinc ion complex.
 KW mating factor alpha 1.
 XX
 XX Homo sapiens.
 XX
 XX Key location/Qualifiers
 FT Protein 1..85
 FT Peptide /Label: mating factor alpha-1
 FT Peptide /Label: A-chain
 FT Peptide /Label: A-chain
 FT Peptide /Label: A-chain
 XX W09507931-A1.
 XX
 XX 23-MAR-1995.
 XX
 XX 16-SEP-1994: 94WD-DK00347.
 XX
 XX 17-SEP-1993: 93DK-0001047.
 PR 02-FEB-1994: 94US-0190823.
 XX
 XX (NOVC) NOVO-NORDISK AS.
 XX
 XX Andersen AS, Halstrom JH, Havelund S, Jonassen J;
 PI Markussen J;
 XX
 XX WPI: 1995-131314/17.
 DR N-PSDB; AAQ86425.
 XX
 XX Acylated insulin deriv. which may be present as a zinc ion
 PT complex - is used to treat diabetes and is rapid acting
 PT
 XX Example 5; Page 78; 100pp; English.
 PS
 XX AAQ86425 encodes AAR71692 mating factor alpha 1-insulin precursor
 CC Arg831. Arg831 comprises the B and A chains of a claimed human
 CC insulin derivative. In the final claimed compsr. they are covalently
 CC connected via disulphide bonds between Cys residues A7/A7 and

CC A20/B19. The derivative, which may be present as a zinc ion
 CC complex, can be used as a fast action treatment for diabetes.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 137 AA;

Query Match 51.5%; Score 302.5; DB 16; Length 137;
 Best Local Similarity 50.0%; Pred. No. 3.6e-18;
 Matches 70; Conservative 4; Mismatches 27; Indels 39; Gaps 4;

QY 2 FPIPLSKLFUNALRAURLHQLAFDTYDEFFAYIPKEO--KYSFLQ-----N 48
 DB 3 FPSI-----FIAVLAASSALAAPVNTTTEDETAQTPAFAVIGYSLZGDFDVAVLPSN 57

QY 49 PLGFG-----PRFVNHLGSHLVEALYLVCGERGFFYPTKTRG 87
 DB 58 STNGLLFINFTIASIAAKEEGVSMARFVNHLGSHLVEALYLVCGERGFFYPTKTRG 117

QY 8% IVEQCCSICSLYLQENYCN 107
 DB 11% IVEQCCSICSLYLQENYCN 137

RESULT 7
 AAR68901
 ID AAR68901 standard; peptide: 56 AA.
 AC AAR68901;
 XX 25-MAR-2003 (updated)
 DT 02-MAR-1995 (first entry)
 XX
 XX Human pro-insulin 3.
 XX
 XX Ito-insulin: A-chain; B-chain; C-chain; disulphide;
 KW mercaptan; chaotropic agent.
 KW
 XX Homo sapiens.
 XX
 XX EP506372-A1.
 XX
 XX 08-JUN-1994.
 XX
 XX 25-NOV-1993: 93EP-0118993.
 XX
 XX 02-DEC-1992: 92DE-4240420.
 XX
 XX (FARH) HCBCHST AG.
 XX
 XX Gerl M, Ludwig J, Obermeier R, Sabel W;
 PI WPI: 1994-177719/22.
 DR
 XX
 XX prodn. of pro-insulin with correct disulphide bridges - by
 PT treating recombinant precursor protein with mercaptan in alkali
 PT and in presence of chaotropic agent, then isolation on
 PT hydrophobic resin
 PT
 XX Disclosure; Page 12; 15pp; German.
 PS
 XX Pro-insulin is produced by treating recombinant precursor protein
 CC with a mercaptan to provide 2-10 SH residues per Cys residue, in
 CC presence of a chaotropic agent and in aq. medium of pH 10-11,
 CC treating the prod. with 3-50 g hydrophobic adsorber resin per 1 aq.
 CC medium of pH 4-7, isolating the adsorbed resin and pro-insulin and
 CC desorbing the pro-insulin. This method produces pro-insulin with
 CC correctly bonded Cys bridges. Compared with known methods it
 CC involves fewer stages (esp. no sulphitolysis or cyanogen bromide
 CC cleavage) and overall losses during purification are reduced, i.e.
 CC the process is quicker and gives better yields.
 CC Sequences of insulin chain A, B and C are given in AAR68895-97.
 CC Sequences of pro-insulin 1-4 are given in AAR68898-901.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 56 AA:
 Query Match 50.9%; Score 299; DB 15; Length 56;
 Best Local Similarity 100.0%; Pred. No. 3.2e-18;
 Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 55 REVNHGCGSHLVEALYLGVGERGFFYTPKTRIGIVEQCCTISCSLYOLENYCN 107
 DB 4 REVNHGCGSHLVEALYLGVGERGFFYTPKTRIGIVEQCCTISCSLYOLENYCN 56

RESULT 8
 AAR78665 standard; peptide: 56 AA.
 XX AAR78665
 AC AAR78665 (first entry)
 DT 03-APR-1996 (first entry)
 XX Proinsulin sequence 3.
 XX Proinsulin: post-translational modification: recombinant production;
 KW protein folding: conformation.
 XX Synthetic.
 XX Key Location/Qualifiers
 FT Region 1..4
 FT /label= R2
 FT /note= "a peptide of 4 amino acids"
 FT Peptide 5..34
 FT /label= R1-(R2-R25) Y
 FT /note= "human insulin B-chain"
 FT Region 35
 FT /label= X
 FT Peptide 36..56
 FT /label= Gly-(A2-A20)-R3
 FT /note= "human insulin A-chain"
 XX EP668292-A2.
 XX 23-AUG-1995.
 XX 09-FEB-1995; 95EP-Q101748.
 XX 18-FEB-1994; 94DE-4405179.
 XX (FARH) HOFCHST AG.
 XX Gerl M, Ludwig J, Obermeier E, Sabel W;
 WPI: 1995-284754/38.
 XX Isolation of insulin that is correctly post translationally
 PT processed - by reacting proinsulin with a mercaptan in the presence
 PT of a chaotropic agent and purification after absorption to hydrophobic
 PT resin
 XX Example 2; Page 13; 15pp; German.
 XX The present sequence is an example of a proinsulin molecule prepared
 CC to the general formula R2-R1-(R2-R20) Y-X-Gly-(A2-A20)-R3 (11). In
 CC formula (11), X = Lys, Arg or a peptide of 2-15 amino acids, prolin,
 CC Lys or Arg at the N- and C-termini; Y = a natural amino acid; R1
 CC phe or a bond; R2 = H, Arg, Lys, a peptide of 2-45 amino acids
 CC contg. Arg or Lys at the N- and C-termini; R3 = a natural amino
 CC acid; (A2-A20) and (B2-B29) are the insulin A- and B-chain sequences
 CC from human or other insulin. The proinsulin molecule (produced in
 CC recombinant E. coli) is reacted with mercaptan at a ratio of 2:10 SH
 CC residues of mercaptan per Cys residue of proinsulin. The reaction
 CC takes place in the presence of a chaotropic auxiliary agent at
 CC pH 10-11 and results in proinsulin with correctly linked cystine

CC Bridges. Reaction with trypsin and opt. carboxypeptidase B yields
 CC correctly folded insulin. The insulin is isolated by absorption on
 CC a hydrophobic resin.
 XX Sequence 56 AA:
 Query Match 50.9%; Score 299; DB 15; Length 56;
 Best Local Similarity 100.0%; Pred. No. 3.2e-18;
 Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 55 REVNHGCGSHLVEALYLGVGERGFFYTPKTRIGIVEQCCTISCSLYOLENYCN 107
 DB 4 REVNHGCGSHLVEALYLGVGERGFFYTPKTRIGIVEQCCTISCSLYOLENYCN 56

RESULT 9
 AAR68899 standard; peptide: 96 AA.
 XX AAR68899
 AC AAR68899;
 XX 25-MAR-2003 (updated)
 DT 02-MAR-1995 (first entry)
 XX Human pro-insulin 2.
 XX Pro-insulin: A-chain; B-chain; C-chain; disulphide;
 KW mercaptan; chaotropic agent.
 XX Homo sapiens.
 XX EP600372-A1.
 XX 08-JUN-1994.
 XX 25-MAY-1993; 93EP-Q1-8993.
 XX 02-DEC-1992; 92DE-4240420.
 XX (FARH) HOFCHST AG.
 XX Gerl M, Ludwig J, Obermeier E, Sabel W;
 WPI: 1994-177718/22.
 XX Prodn. of pro-insulin with correct disulphide bridges - by
 PT treating recombinant precursor protein with mercaptan in alkali
 PT and in presence of chaotropic agent, then isolation on
 PT hydrophobic resin
 XX Disclosure: Page 11; 15pp; German.
 XX Pro-insulin is produced by treating recombinant precursor protein
 CC with a mercaptan to provide 2:10 SH residues per Cys residue. In
 CC presence of a chaotropic agent and in aq. medium of pH 10-11,
 CC treating the prod. with 3-50 g hydrophobic adsorber resin per 1 aq.
 CC medium of pH 4-7, isolating the adsorbed resin and pro-insulin and
 CC desorbing the pro-insulin. This method produces pro-insulin with
 CC correctly bonded Cys bridges. Compared with known methods it
 CC involves fewer stages (esp. no sulphitolysis or cyanogen bromide
 CC cleavage) and overall losses during purification are reduced, i.e.
 CC the process is quicker and gives better yields.
 CC Sequences of insulin chain A, B and C are given in AAR68895-97.
 CC Sequences of pro-insulin 1-4 are given in AAR68898-901.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX Sequence 96 AA:
 Query Match 50.9%; Score 299; DB 15; Length 96;
 Best Local Similarity 100.0%; Pred. No. 5.1e-18;
 Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 55 REVNHGCGSHLVEALYLGVGERGFFYTPKTRIGIVEQCCTISCSLYOLENYCN 107

DB 44 RFVNHLCGSHLVEALYVGCGRGFFYPKTRGIVECCCTSIQSYOLENYCN 96

RESULT 10

AAR78662
ID AAR78662 standard; protein: 96 AA.

XX AAR78662;

AC 03-APR-1996 (first entry)

XX Fusion protein contg. proinsulin sequence 3.

XX Proinsulin; post-translational modification: recombinant production;
KW protein folding; conformation.

XX Synthetic.

XX Key Location/Qualifiers

XX Region 41..44

XX Peptide /label= R2

XX /note= "a peptide of 4 amino acids"

XX Region 45..74

XX Peptide /label= R1-(B2-B29)-Y

XX /note= "human insulin B-chain"

XX Region 75

XX Peptide /label= X

XX /note= "human insulin A-chain"

XX Region 75..96

XX Peptide /label= R1-(A2-A29)-K3

XX /note= "human insulin A-chain"

XX EP666232-A2.

XX 23-MAR-1995.

XX 09-FEB-1995; 95EP-0201748

XX 18-FEB-1994; 94DE-4405179

XX (FARR) HOECHST AG.

XX Gerl M, Ludwig J, Oberwieser R, Sabel W;

XX WPI: 1995-284754/38.

XX Isolation of insulin that is correctly post-transcriptionally
XX processed - by reacting pro-insulin with a mercaptan in the presence
XX of a chaotropic agent and purification after absorption to hydrophobic
XX resin.

XX Example 2; Page 8; 16pp; German.

XX The present sequence is that of a fusion protein, produced by
XX which contains an example of a proinsulin molecule corresp.
XX to the general formula R2-R1-(B2-B29)-Y-X-Gly-(A2-A20)-R3 (11). In
XX formula (11); X = Lys, Arg or a peptide of 2-35 amino acids contg.
XX Lys or Arg at the N- and C-termini; Y = a natural amino acid; R1 =
XX Phe or a bond; R2 = H, Arg, Lys, a peptide of 2-45 amino acids
XX contg. Arg or Lys at the N- and C-termini; R3 = a natural amino
XX acid; (A2-A20) and (B2-B29) are the insulin A- and B-chain sequences
XX from human or other insulin. The proinsulin molecule, released by
XX cyanogen bromide, is reacted with mercaptan at a ratio of 2:10 SH
XX residues of mercaptan per Cys residue of proinsulin. The reaction
XX takes place in the presence of a chaotropic auxiliary agent at
XX pH 10-11 and results in proinsulin with correctly linked cystine
XX bridges. Reaction with trypsin and opt. carboxypeptidase B yields
XX correctly folded insulin. The insulin is isolated by absorption on
XX a hydrophobic resin.

XX Sequence 96 AA;

XX Query Match 50.9%; Score 299; DB 16; Length 96;

Best Local Similarity 100.0%; Pred. No. 5,1e-18;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC 55 RFVNHLCGSHLVEALYVGCGRGFFYPKTRGIVECCCTSIQSYOLENYCN 107

DB 44 RFVNHLCGSHLVEALYVGCGRGFFYPKTRGIVECCCTSIQSYOLENYCN 96

RESULT 11

AAR71694
ID AAR71694 standard; Protein: 145 AA.

XX AAR71694;

XX 25-MAR-2003 (updated)

XX 20-NOV-1995 (first entry)

XX Mating factor alpha 1-Insulin precursor ArgB1, ArgB31 N-terminal.

XX Human insulin precursor ArgB1, ArgB31; diabetes; zinc ion complex;
KW mating factor alpha 1; N-terminal EEAEAEAR.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Protein 1..85

XX Peptide /label= mating factor alpha-1

XX /label= N-terminal peptide

XX Peptide 94..124

XX /label= B-chain

XX /label= A-chain

XX W09507937-A1.

XX 23-MAR-1995.

XX 16-SEP-1994; 94WO-DK0347.

XX 17-SEP-1993; 93DK-0001044.

XX 02-FEB-1994; 94US-0190829.

XX (NOVO) NOVO-N-BDISK AS.

XX Anderson AS, Halstrom JB, Havelund S, Jørgensen I;

XX Markussen J;

XX WPI: 1995-131314/17.

XX N-PSD3; AA086429.

XX Acylated insulin deriv. which may be present as a zinc ion
XX complex - is used to treat diabetes and is rapid acting.
XX Example 5; Pages 82-83; 100pp; English.

XX AA086429 encodes AAR71694 mating factor alpha 1-Insulin precursor
XX ArgB1, ArgB31 N-terminal EEAEAEAR. The insulin precursor comprises
XX the H and A chains of a claimed human insulin derivative preceded
XX by the N-terminal amino acids EEAEAEAR. In the final claimed compsn.
XX they are covalently connected via disulphide bonds between Cys
XX residues A7/B7 and A20/B19. The derivative, which may be present
XX as a zinc ion complex, can be used as a fast action treatment for
XX diabetes.
XX (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 145 AA;

XX Query Match 50.9%; Score 299; DB 16; Length 145;

Best Local Similarity 100.0%; Pred. No. 7,3e-18;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC 55 RFVNHLCGSHLVEALYVGCGRGFFYPKTRGIVECCCTSIQSYOLENYCN 107

DB 93 RVVNOHLCGSHLVREALYLVGGERGFYTPKTRGIVEQCCTSCSYOLENYCN 146

RESULT 12

AA71695

ID AAR71695 standard; protein; 146 AA.

XX AC AAR71695;

XX DT 25-MAR-2003 (updated)

XX DT 20-NOV-1995 (first entry)

XX DE Mating factor alpha 1-insulin precursor ArgB1, ArgB3; diabetes; zinc ion complex

XX KW Human insulin precursor ArgB1, ArgB3; diabetes; zinc ion complex

XX KW mating factor alpha 1; N-terminal EEAAAEAFR.

XX OS Homo sapiens.

XX PH Key location/Qualifiers

FT Protein: 1..85

FT Peptide /label: mating factor alpha-1

FT Peptide 86..94

FT Peptide /label: N-terminal peptide

FT Peptide 95..125

FT Peptide /label: B-chain

FT Peptide 126..146

FT Peptide /label: A-chain

XX W09507931-A1.

XX PN 23-MAR-1995.

XX PD 16-SEP-1994; 94WO-DK00347.

XX PR 17-SEP-1993; 93DK-0001044.

XX PR 02-FEB-1994; 94CS-0190829.

XX PA (NOVO) NOVO-NORDISK AS.

XX PI Andersen AS, Halstrom JB, Havemand S, Jonassen I;

XX PI Markussen J;

XX DR WPI: 1995-131114/17.

XX DR N-PSDB: AA086432.

XX PT Acylated insulin deriv. which may be present as a zinc ion complex - is used to treat diabetes and is rapid acting.

XX PS Example 6; Page 85; 100pp; English.

XX CC AA086432 encodes AAR71695 mating factor alpha 1-insulin precursor ArgB1, ArgB3 N-terminal EEAAAEAFR. The insulin precursor comprises the B and A chains of a claimed human insulin derivative prepared by the N-terminal amino acids EEAAAEAFR. In the final claimed complex they are covalently connected via disulphide bonds between Cys residues A7/B7 and A26/H19. The derivative, which may be present as a zinc ion complex, can be used as a fast action treatment for diabetes.

XX CC (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 146 AA;

Query Match: 50.9%; Score 254; DB 16; Length 146;

Best Local Similarity: 100.0%; Pred. No. 7,44-36;

Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 13

AA71695

ID AAY42859 standard; protein; 52 AA.

XX AC AAY42859;

XX DT 19-JAN-2000 (first entry)

XX DE Human insulin precursor, SEQ ID 5.

XX KW Insulin precursor; growth hormone; chaperone; intramolecular; folding; conformation; chimeric protein; cleavable; recombinant; production; yield.

XX OS Homo sapiens

XX EN W09503002-A1.

XX ID 07-OCT-1999.

XX PF 31-MAR-1998; 98WO-CN00052.

XX PR 31-MAR-1998; 98WO-CN00052.

XX PA (TONG) TONGHUA GANTECH BIOTECHNOLOGY LTD.

XX PT Gan Z;

XX DR WPI: 1999-610839/52.

XX PT New chimeric proteins containing human growth hormone fragment, used particularly for the production of human insulin

XX PS Claim 12; Page 29-30; 46pp; English.

XX CC This sequence represents a human insulin precursor comprising insulin A and B chains. This insulin precursor is a component of the chimeric proteins hGH-mini-proinsulin (AAY42860) and the chimeric protein given in AAY42861. These chimeric proteins additionally contain an N-terminal fragment of human growth hormone (hGH) and a cleavable peptide linker (AAY42857). The hGH portion of the chimeric protein acts as an intramolecular chaperone (IMC) for the insulin precursor, enabling it to fold correctly. The cleavable peptide linker has a C-terminal Arg residue which enables the hGH portion of the chimeric protein to be removed after folding has taken place. Production of recombinant human insulin via an hGH-proinsulin chimeric protein can provide human insulin with correctly linked cysteine bridges with fewer necessary procedural steps, and hence resulting in a higher yield of human insulin. The IMC sequences not only protect insulin sequences from intracellular degradation by a microorganism host, but also promote the folding of the fused insulin precursor, facilitate the solubility of the fusion protein and decrease the intermolecular interactions among the fusion proteins, thus allowing folding of the fused insulin precursor at commercially useful high concentrations. The procedural steps of cyanogen bromide cleavage, oxidative aliphaticolysis and related purification steps can thus be eliminated, along with the use of high concentrations of mercaptan or the use of hydrophobic absorbent resins.

XX SQ Sequence 52 AA;

Query Match: 50.1%; Score 294; DB 20; Length 52;

Best Local Similarity: 100.0%; Pred. No. 7,4e-18;

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 56 FVNOHLCGSHLVREALYLVGGERGFYTPKTRGIVEQCCTSCSYOLENYCN 107

DB 1 FVNOHLCGSHLVREALYLVGGERGFYTPKTRGIVEQCCTSCSYOLENYCN 52

RFSD0114

AA71695

ID AAR04582 standard; protein; 57 AA.

XX

GenDate version 5.1.6
Copyright (c) 1992 - 2003 Caparon Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 12:37:10 Search time 13.7 seconds
(without alignments)
265,217 Million cell updates/sec

Title: US-09-423-100-6

Perfect score: 587

Sequence: 1 NP01P1SK3.FDNAM.BARR.....IVGQNTSLPSQIDNVYR 107

Scoring table: BL2SUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310856 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.ppt:*
2: /cgn2_6/ptodata/1/1aa/5A_COMB.ppt:*
3: /cgn2_6/ptodata/1/1aa/5A_COMB.ppt:*
4: /cgn2_6/ptodata/1/1aa/5A_COMB.ppt:*
5: /cgn2_6/ptodata/1/1aa/PC1US_COMB.ppt:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.ppt:*

Pred. No. is the number of results predicted by Gmatch to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	304	51.8	63	1	US-08-160-376A-6
2	302.5	51.5	137	3	US-08-400-256-15
3	302.5	51.5	137	3	US-08-975-365-15
4	299.5	51.0	66	1	US-08-293-060P-6
5	299	50.9	56	1	US-08-160-376A-6
6	299	50.9	56	1	US-08-389-487-11
7	299	50.9	96	1	US-08-160-376A-6
8	299	50.9	96	1	US-08-389-487-11
9	299	50.9	145	3	US-08-400-256-15
10	299	50.9	145	3	US-08-975-365-15
11	299	50.9	146	1	US-08-400-256-15
12	299	50.9	146	3	US-08-975-365-15
13	293	49.3	57	1	US-08-330-733A-44
14	287	48.9	65	3	US-08-900-574-5
15	286.5	48.6	66	3	US-08-900-574-5
16	286	48.7	67	3	US-08-900-574-5
17	284.5	48.5	65	1	US-08-468-674P-71
18	284.5	48.5	65	1	US-08-780-571-71
19	284.5	48.5	124	3	US-08-012-669F-16
20	284	48.4	138	3	US-08-932-582-19
21	284	48.4	140	1	US-08-400-256-15
22	284	48.4	140	1	US-08-400-256-15
23	284	48.4	140	3	US-08-975-365-15
24	284	48.4	140	3	US-08-975-365-15
25	283.5	48.3	53	1	US-08-233-617-4
26	283.5	48.3	53	4	US-08-981-988A-42
27	283.5	48.3	117	3	US-09-012-669F-37

US-08-160-376A-6
Sequence 6, Application US/08160376A
Patent No. 5473049
GENERAL INFORMATION:
APPLICANT: Obermeier, Ranier
APPLICANT: Gerl, Martin
APPLICANT: Ludwig, Jurgen
APPLICANT: Sabel, Walter
TITLE OF INVENTION: Process For Obtaining Protein
TITLE OF INVENTION: Possessing Correctly Linked
TITLE OF INVENTION: Cystine Bridges
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenneth A. Genoni, Esq.
STREET: Rt. 202-206 No. 5473049th/P.O. Box 2500
CITY: Somerville
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 08876-1258
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM 386
OPERATING SYSTEM: WINDOWS 3.1
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/160.376A
FILING DATE: December 1, 1993
CLASSIFICATION: 530
APPLICATION DATA:
APPLICATION NUMBER: GE P 4240420.7
FILING DATE: December 2, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Barbara V. Maurer, Esq.
REGISTRATION NUMBER: 31,287
REFERENCE/DOCKET NUMBER: ROE 92/P 384
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 231-4079
TELEFAX: (908) 231-2255
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 63 Amino Acids
TYPE: Amino Acid (AA)
TOPOLOGY: not relevant
US-08-160-376A-6

ALIGNMENTS

Sequence 6, Application US/08160376A
Patent No. 5473049
GENERAL INFORMATION:
APPLICANT: Obermeier, Ranier
APPLICANT: Gerl, Martin
APPLICANT: Ludwig, Jurgen
APPLICANT: Sabel, Walter
TITLE OF INVENTION: Process For Obtaining Protein
TITLE OF INVENTION: Possessing Correctly Linked
TITLE OF INVENTION: Cystine Bridges
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenneth A. Genoni, Esq.
STREET: Rt. 202-206 No. 5473049th/P.O. Box 2500
CITY: Somerville
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 08876-1258
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM 386
OPERATING SYSTEM: WINDOWS 3.1
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/160.376A
FILING DATE: December 1, 1993
CLASSIFICATION: 530
APPLICATION DATA:
APPLICATION NUMBER: GE P 4240420.7
FILING DATE: December 2, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Barbara V. Maurer, Esq.
REGISTRATION NUMBER: 31,287
REFERENCE/DOCKET NUMBER: ROE 92/P 384
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 231-4079
TELEFAX: (908) 231-2255
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 63 Amino Acids
TYPE: Amino Acid (AA)
TOPOLOGY: not relevant
US-08-160-376A-6

Query Match 53.8%; Score 304; DB 1; Length 63;
Best Local Similarity 94.7%; Pred. No. 7.3e-29;
Matches 54; Conservative 0; Mismatches 3; Indels 0; Gaps 0;


```

1 COUNTRY: USA
2 ZIP: 20005-3315
3 COMPUTER READABLE FORM:
4 MEDIUM TYPE: Floppy disk
5 OPERATING SYSTEM: PC-DOS/MS-DOS
6 SOFTWARE: Patent Release #1.0, Version #1.0
7 CURRENT APPLICATION DATA:
8 FILING DATE: 08-AUG-1994
9 CLASSIFICATION: 435
10 ATTORNEY/AGENT INFORMATION:
11 NAME: Binardi, Carol P.
12 REGISTRATION NUMBER: 32,220
13 REFERENCE/DOCKET NUMBER: 02481-1105-0200
14 TELECOMMUNICATION INFORMATION:
15 TELEPHONE: (202) 408-4363
16 TELEFAX: (202) 408-4400
17 INFORMATION FOR SEQ ID NO: 5:
18 SEQUENCE CHARACTERISTICS:
19 LENGTH: 66 amino acids
20 TYPE: amino acid
21 STRANDEDNESS: single
22 TOPOLOGY: linear
23 MOLECULE TYPE: peptide
24 US-08-291-0603-5
25
26 Query Match 51.0%; Score 299.5; DB 1; Length 66;
27 Best Local Similarity 91.7%; Pred. No. 2.5e-26;
28 Matches 55; Conservative 1; Mismatches 3; Indels 1; Gaps 1;
29
30 QY 4B NPLGTPFVNHLCGSHVEALYLVCGERGFFYTPKTRGIVEOCCTSGSLYLENYCN 107
31 DB 8 DPNSNG-RVNHQLCCHLVEALYLVCGERGFFYTPKTRGIVEOCCTSGSLYLENYCN 66
32
33 RESULT 5
34 US-08-160-376A-7
35 Sequence 1, Application US/0810376A
36 Patent No. 5473049
37 GENERAL INFORMATION:
38 APPLICANT: Obermeier, Rainer
39 APPLICANT: Geil, Martin
40 APPLICANT: Ludwig, Jürgen
41 APPLICANT: Sabel, Walter
42 TITLE OF INVENTION: Process for Obtaining Proteinulin
43 TITLE OF INVENTION: Possessing Correctly Linked
44 TITLE OF INVENTION: Cysteine Bridges
45 NUMBER OF SEQUENCES: 7
46 CORRESPONDENCE ADDRESS:
47 ADDRESSEE: Kenneth A. Gibson, Esq.
48 STREET: Rt. 202-206 No. 547049th/P.O. Box 4509
49 CITY: Somerville
50 STATE: New Jersey
51 COUNTRY: U.S.A.
52 ZIP: 08876-1258
53 COMPUTER READABLE FORM:
54 MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
55 OPERATING SYSTEM: WINDOWS 3.1
56 SOFTWARE: WORDPERFECT 5.1
57 CURRENT APPLICATION DATA:
58 FILING DATE: December 1, 1993
59 CLASSIFICATION: 530
60 PRIOR APPLICATION DATA:
61 APPLICATION NUMBER: SE P 4240420.7
62 FILING DATE: December 2, 1992
63 ATTORNEY/AGENT INFORMATION:
64 NAME: Barbara V. Maurer, Esq.
65 REGISTRATION NUMBER: 31,287
66 REFERENCE/DOCKET NUMBER: HOE 92/F 384
67 TELECOMMUNICATION INFORMATION:

```

```

1 TELEPHONE: (908) 231-4079
2 TELEFAX: (908) 231-2255
3 INFORMATION FOR SEQ ID NO: 7:
4 SEQUENCE CHARACTERISTICS:
5 LENGTH: 56 Amino Acids
6 TYPE: Amino Acid (AA)
7 TOPOLOGY: not relevant
8 US-08-160-376A-7
9
10 Query Match 50.9%; Score 299; DB 1; Length 56;
11 Best Local Similarity 100.0%; Pred. No. 2.4e-28;
12 Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
13
14 QY 55 RFVNHLCGSHLVEALYLVCGERGFFYTPKTRGIVEOCCTSGSLYLENYCN 107
15 DB 4 RFVNHLCGSHLVEALYLVCGERGFFYTPKTRGIVEOCCTSGSLYLENYCN 56
16
17 RESULT 6
18 US-08-349-487-11
19 Sequence 11, Application US/08389487
20 Patent No. 5663251
21 GENERAL INFORMATION:
22 APPLICANT: Obermeier, Rainer
23 APPLICANT: Geil, Martin
24 APPLICANT: Ludwig, Jürgen
25 APPLICANT: Sabel, Walter
26 TITLE OF INVENTION: Process for Obtaining Insulin Having
27 TITLE OF INVENTION: Correctly Linked Cysteine Bridges
28 NUMBER OF SEQUENCES: 12
29 CORRESPONDENCE ADDRESS:
30 ADDRESSEE: Dunner
31 STREET: 1300 I Street, N.W.
32 CITY: Washington
33 STATE: D.C.
34 COUNTRY: United States of America
35 ZIP: 20005-3315
36 COMPUTER READABLE FORM:
37 MEDIUM TYPE: Floppy disk
38 COMPUTER: IBM PC compatible
39 OPERATING SYSTEM: PC-DOS/MS-DOS
40 SOFTWARE: Patent Release #1.0, Version #1.25
41 CURRENT APPLICATION DATA:
42 APPLICATION NUMBER: US/08/389,487
43 FILING DATE:
44 CLASSIFICATION: 530
45 ATTORNEY/AGENT INFORMATION:
46 NAME: Binardi, Carol P.
47 REGISTRATION NUMBER: 32,220
48 REFERENCE/DOCKET NUMBER: 02481-1424-0000
49 TELECOMMUNICATION INFORMATION:
50 TELEPHONE: 202-408-4000
51 TELEFAX: 202-408-4400
52 INFORMATION FOR SEQ ID NO: 11:
53 SEQUENCE CHARACTERISTICS:
54 LENGTH: 56 amino acids
55 TYPE: amino acid
56 STRANDEDNESS: single
57 TOPOLOGY: linear
58 MOLECULE TYPE: peptide
59 US-08-369-487-11
60
61 Query Match 50.9%; Score 299; DB 1; Length 56;
62 Best Local Similarity 100.0%; Pred. No. 2.4e-28;
63 Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
64
65 QY 55 RFVNHLCGSHLVEALYLVCGERGFFYTPKTRGIVEOCCTSGSLYLENYCN 107
66 DB 4 RFVNHLCGSHLVEALYLVCGERGFFYTPKTRGIVEOCCTSGSLYLENYCN 56
67
68 RESULT 7

```

```
US-08-160-376A-5
: Sequence 5, Application US/0816-376A
: Patent No. 5473049
: GENERAL INFORMATION:
: APPLICANT: Obermeier, Rainer
: APPLICANT: Gerl, Martin
: APPLICANT: Ludwig, Jurgen
: APPLICANT: Sabell, Walter
: TITLE OF INVENTION: Process For Obtaining Proinsulin
: TITLE OF INVENTION: Possessing Correctly Linked
: TITLE OF INVENTION: Cystine Bridges
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Kenneth A. Genovai, Esq.
: STREET: Rt. 202-206 No. 5473049L/P.O. Box 2500
: CITY: Somerville
: STATE: New Jersey
: COUNTRY: U.S.A.
: ZIP: 08876-1258
: COMPUTER READABLE FORM:
: MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
: COMPUTER: IBM 386
: OPERATING SYSTEM: WINDOWS 3.1
: SOFTWARE: WORDPERFECT 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/160.375A
: FILING DATE: December 1, 1993
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GE P 4240420.7
: FILING DATE: December 2, 1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Barbara V. Maurel, Esq.
: REGISTRATION NUMBER: 31,287
: REFERENCE/DOCKET NUMBER: HCE 92/P 384
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (908) 231-4575
: TELEFAX: (908) 231-2255
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 96 Amino Acids
: TYPE: Amino Acid (AA)
: TOPOLOGY: Not relevant
: US-08-160-376A-5

Query Match 50.9% Score 299; DB 1; Length 96;
Best Local Similarity 100.0%; Prod. No. 4.8e-28;
Matches 53; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 55 RYVNOHCGSHLVLYVCGERGFFYTPKTRGIVVEQCTTCSLYQLENYCN 107
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 44 RYVNOHCGSHLVLYVCGERGFFYTPKTRGIVVEQCTTCSLYQLENYCN 96

RESULT 8
US-08-389-487-8
: Sequence 8, Application US/0816-9487
: Patent No. 5663291
: GENERAL INFORMATION:
: APPLICANT: Obermeier, Rainer
: APPLICANT: Gerl, Martin
: APPLICANT: Ludwig, Jurgen
: APPLICANT: Sabell, Walter
: TITLE OF INVENTION: Process for Obtaining Insulin Having
: TITLE OF INVENTION: Correctly Linked Cystine Bridges
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESSES: Finnegan, Henderson, Farabow, Garrett &
: ADDRESSEE: Dunner
: STREET: 1300 I Street, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: United States of America

US-08-160-376A-5
Query Match 50.9% Score 299; DB 1; Length 96;
Best Local Similarity 100.0%; Prod. No. 4.8e-28;
Matches 53; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 55 RYVNOHCGSHLVLYVCGERGFFYTPKTRGIVVEQCTTCSLYQLENYCN 107
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 44 RYVNOHCGSHLVLYVCGERGFFYTPKTRGIVVEQCTTCSLYQLENYCN 96

RESULT 9
US-08-400-256-45
: Sequence 45, Application US/08400256
: Patent No. 5750497
: GENERAL INFORMATION:
: APPLICANT: Havelund, Svend
: APPLICANT: Halstrom, John
: APPLICANT: Conassen, Ib
: APPLICANT: Andersen, Asger Sloth
: APPLICANT: Markussen, Jan
: TITLE OF INVENTION: ACYLATED INSULIN
: NUMBER OF SEQUENCES: 49
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: No. 5750497e No. 5750497disk of No. 5750497th America, Inc.
: STREET: 405 Lexington Avenue, 54th Floor
: CITY: New York
: STATE: New York
: COUNTRY: United States of America
: ZIP: 10174-6401
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/400,256
: FILING DATE: 03-MAR-1995
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Lambiris, Elias J.
: REGISTRATION NUMBER: 33,728
: REFERENCE/DOCKET NUMBER: 3985.220-US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-867-0123
: TELEFAX: 212-878-9655
: INFORMATION FOR SEQ ID NO: 45:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 145 amino acids
```



```

: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-400-256-45

Query Match      50.9%  Score 299; DB 1; Length 146;
Best Local Similarity 100.0%; Pred. No. 80-28;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 55 RFVNHGCGSHLVREALYVCGRGFFYPKTRGIVEQCCTSCSLYQLENYCN 107
DB 93 RFVNHGCGSHLVREALYVCGRGFFYPKTRGIVEQCCTSCSLYQLENYCN 146

RESULT 10
US-08-975-365-45
: Sequence 45, Application US/08975365
: Patent No. 6011007
: GENERAL INFORMATION:
: APPLICANT: Havelund, Svend
: APPLICANT: Halstrom, John
: APPLICANT: Jonassen, Ib
: APPLICANT: Andersen, Asger Sloth
: TITLE OF INVENTION: ACYLATED INSULIN
: NUMBER OF SEQUENCES: 45
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: No. 6011007 No. 6011007disk of No. 6011007th America, Inc.
: STREET: 405 Lexington Avenue, 64th Floor
: CITY: New York
: STATE: New York
: COUNTRY: United States of America
: ZIP: 10174-6401
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/975,365
: FILING DATE: 03-MAR-1995
: CLASSIFICATION: 5:4
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/400,256
: FILING DATE: 03-MAR-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Lambiris, Elias J.
: REGISTRATION NUMBER: 33,728
: REFERENCE/DOCKET NUMBER: 3985,220-CS
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-867-0123
: TELEFAX: 212-878-9655
: INFORMATION FOR SEQ ID NO: 45:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 146 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-975-365-45

Query Match      50.9%  Score 299; DB 1; Length 146;
Best Local Similarity 100.0%; Pred. No. 80-28;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 55 RFVNHGCGSHLVREALYVCGRGFFYPKTRGIVEQCCTSCSLYQLENYCN 107
DB 93 RFVNHGCGSHLVREALYVCGRGFFYPKTRGIVEQCCTSCSLYQLENYCN 146

RESULT 11
US-08-400-256-4B
: Sequence 48, Application US/08400256
: Patent No. 5750497

```

```

: GENERAL INFORMATION:
: APPLICANT: Havelund, Svend
: APPLICANT: Halstrom, John
: APPLICANT: Jonassen, Ib
: APPLICANT: Andersen, Asger Sloth
: TITLE OF INVENTION: ACYLATED INSULIN
: NUMBER OF SEQUENCES: 49
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: No. 5750497 No. 5750497disk of No. 5750497th America, Inc.
: STREET: 405 Lexington Avenue, 64th Floor
: CITY: New York
: STATE: New York
: COUNTRY: United States of America
: ZIP: 10174-6401
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/400,256
: FILING DATE: 03-MAR-1995
: CLASSIFICATION: 5:4
: ATTORNEY/AGENT INFORMATION:
: NAME: Lambiris, Elias J.
: REGISTRATION NUMBER: 33,728
: REFERENCE/DOCKET NUMBER: 3985,220-US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-867-0123
: TELEFAX: 212-878-9655
: INFORMATION FOR SEQ ID NO: 48:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 146 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-400-256-48

Query Match      50.9%  Score 299; DB 1; Length 146;
Best Local Similarity 100.0%; Pred. No. 80-28;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 55 RFVNHGCGSHLVREALYVCGRGFFYPKTRGIVEQCCTSCSLYQLENYCN 107
DB 94 RFVNHGCGSHLVREALYVCGRGFFYPKTRGIVEQCCTSCSLYQLENYCN 146

RESULT 12
US-08-975-365-4B
: Sequence 48, Application US/08975365
: Patent No. 6011007
: GENERAL INFORMATION:
: APPLICANT: Havelund, Svend
: APPLICANT: Halstrom, John
: APPLICANT: Jonassen, Ib
: APPLICANT: Andersen, Asger Sloth
: TITLE OF INVENTION: ACYLATED INSULIN
: NUMBER OF SEQUENCES: 49
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: No. 6011007 No. 6011007disk of No. 6011007th America, Inc.
: STREET: 405 Lexington Avenue, 64th Floor
: CITY: New York
: STATE: New York
: COUNTRY: United States of America
: ZIP: 10174-6401
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:

```


GenCore version: 5.1.5
Copyright (c) 1993 - 2003 Computer Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 12:38:55 : Search time 33.7237 seconds
(without alignments)
472.415 Million cell operations/sec

Title: US-09-423-100-6

Perfect score: 587

Sequence: 1 MFETIP:SLPLDNLMLRAIRK.....INRGQNTDLSLYLENVKK (2)

Scoring table: BLASTSUM62

Gapop 13.0 , Gapext 2.5

Searched: 550269 seqs, 148493353 residues

Total number of hits satisfying chosen parameters: 550269

Minimum DB seq length: 6

Maximum DB seq length: 200000000

Post-processing: Minimum Match: 0%

Listing first 45 summaries

Database :

Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	597	100.0	107	14	US-10-054-873-6
2	555.5	94.6	150	14	US-10-054-873-7
3	294	50.1	52	14	US-10-054-873-5
4	264.5	48.5	124	10	US-09-894-711-5
5	284	48.4	138	10	US-09-894-711-4
6	278.5	47.4	51	13	US-09-854-935R-5
7	278.5	47.4	51	14	US-10-028-410-3
8	278	47.4	127	9	US-09-280-030-6
9	277	47.2	96	10	US-09-947-554-4
10	275.5	46.9	124	9	US-09-736-611-12
11	275.5	46.9	124	9	US-09-740-353-12
12	275.5	46.9	124	10	US-09-894-711-2
13	275.5	46.9	125	9	US-09-736-611-10
14	275.5	46.9	125	9	US-09-740-353-10
15	275.5	46.9	125	10	US-09-894-711-10

275.5 46.9 147 9 US-09-736-611-8
275.5 46.9 147 9 US-09-740-353-7
274 46.7 144 9 US-09-736-611-6
274 46.7 144 9 US-09-740-353-5
274 46.7 146 10 US-09-894-711-5
273 46.5 50 14 US-10-056-009A-3
271 46.2 96 10 US-09-947-563-5
269.5 45.9 130 9 US-09-280-030-62
267 45.5 86 10 US-09-978-380-1
267 45.5 86 11 US-09-958-935B-4
267 45.5 86 14 US-10-028-410-2
267 45.5 86 14 US-10-054-873-4
267 45.5 110 9 US-09-285-656-125
267 45.5 110 9 US-09-815-229-3
267 45.5 110 10 US-09-504-409A-9
267 45.5 110 12 US-09-969-749C-6
267 45.5 110 15 US-10-328-813-2
267 45.5 110 15 US-10-328-813-2
260 44.3 43 14 US-10-054-873-1
260 44.3 92 14 US-10-054-873-2
255.5 42.5 131 11 US-09-984-010-23
255.5 42.5 131 12 US-10-153-207-1
255.5 42.5 191 12 US-10-400-377-1
255.5 42.5 191 12 US-10-400-708-1
255.5 42.5 191 12 US-10-398-148-1
255.5 42.5 214 12 US-10-153-207-6
255.5 42.5 217 9 US-09-929-918-5
255.5 42.5 245 9 US-09-280-030-56
255.5 42.5 144 10 US-09-894-711-7
249.5 42.5 217 9 US-09-853-688-2

ALIGNMENTS

RESULT :
US-10-054-873-6
Sequence 6, Application US/10054873
Publication No. US20020164712A1
GENERAL INFORMATION:
APPLICANT: Gan, Zhong Ru
TITLE OF INVENTION: Chimeric Protein Containing an Intramolecular Chaperone-Like Sequence
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMETER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/054,873
FILING DATE: 22-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WC PCT/CN98/000052
FILING DATE: 31-MAR-1998
APPLICATION NUMBER: US 09/423,100
FILING DATE: 11-DEC-2000
ATTORNEY/AGENT INFORMATION:
NAME: Mycroft, Frank J
REGISTRATION NUMBER: 46,946
REFERENCE/DOCKET NUMBER: 020167-000313005
INFORMATION FOR SEQ. IT. NO. 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid

```

: STRANDEDNESS: <Unknown>
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-054-873-6

Query Match
Best Local Similarity 100.0%; Score 587; DB 14; Length 107;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MFPTIPSLRFDNAMLRAIRLHOLAFDTYGEFEFAYPKCKKYSFLGNPLGTGPRFVNCH 60
DB 1 MFPTIPSLRFDNAMLRAIRLHOLAFDTYGEFEFAYPKCKKYSFLGNPLGTGPRFVNCH 60
OY 61 LCGSHLVEALYVGCGRGGFFYTPKTRGIVGQCCTSIICSLYLENYCN 107
DB 61 LCGSHLVEALYVGCGRGGFFYTPKTRGIVGQCCTSIICSLYLENYCN 107

RESULT 2
US-10-054-873-7
: Sequence 7, Application US/10054873
: Publication No. US20020164712A1
: GENERAL INFORMATION:
: APPLICANT: Gan, Zhong Ru
: TITLE OF INVENTION: Chimeric Protein Containing an
: Intramolecular Chapterone-Like Sequence
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Townsend and Crew LLP
: STREET: Two Embarcadero Center, Eighth Floor
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-3834
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/10/054,873
: FILING DATE: 22-Jan-2002
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: WO PCT/CN98/00052
: FILING DATE: 31-MAR-1998
: APPLICATION NUMBER: US 09/423,100
: FILING DATE: 11-DEC-2000
: ATTORNEY/AGENT INFORMATION:
: NAME: Mycroft, Frank J
: REGISTRATION NUMBER: 46,946
: REFERENCE/DOCKET NUMBER: 020167-00013005
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 150 amino acids
: TYPE: amino acid
: STRANDEDNESS: <Unknown>
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-054-873-7

Query Match
Best Local Similarity 94.6%; Score 555.5; DB 14; Length 150;
Matches 107; Conservative 0; Mismatches 5; Indels 43; Gaps 1;

OY 1 MFPTIPSLRFDNAMLRAIRLHOLAFDTYGEFEFAYPKCKKYSFLGNP----- 49
DB 1 MFPTIPSLRFDNAMLRAIRLHOLAFDTYGEFEFAYPKCKKYSFLGNPLTSFSSSL 60
OY 50 -----LGIQPRFVNHLCGSHLVEALYVCGER 77
DB 50 -----LGIQPRFVNHLCGSHLVEALYVCGER 77

```



```

Query Match      47.4%  Score 275.39  DB 14  Length 517
Best Local Similarity 98.1%  Pred. No. 2,90-267
Matches 51: Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 56 FVNQHLGSHLVEALYLVCGERGFFYTPKTRKGVVEQVQCTSLVQLENYCN 107
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRKGVVEQVQCTSLVQLENYCN 51

RESULT 8
US-09-280-030-63
: Sequence 63, Application US/09280030A
: Patent No. US20010021515A1
: GENERAL INFORMATION:
: APPLICANT: Sato, Seiji
: APPLICANT: Higashikuni, Naohiko
: APPLICANT: Kudo, Toshiyuki
: APPLICANT: Kondo, Masaki
: TITLE OF INVENTION: DNAS ENCODING NEW FUSION PROTEINS AND PROCESSORS FOR
: TITLE OF INVENTION: PREPARING USEFUL POLYPEPTIDES THROUGH EXPRESSION OF THE
: TITLE OF INVENTION: DNAS
: FILE REFERENCE: 382,1226
: CURRENT APPLICATION NUMBER: US/05/280,030A
: CURRENT FILING DATE: 1999-03-26
: EARLIER APPLICATION NUMBER: JP10-87339/1998
: EARLIER FILING DATE: 1998-03-31
: NUMBER OF SEQ ID NOS: 56
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 63
: LENGTH: 117
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: Designated as
: OTHER INFORMATION: as amino acid sequence of
: OTHER INFORMATION: MWSP-MWPE10-Ket-Pro-insulin
US-09-280-030-63

Query Match      47.4%  Score 274;  DB 3;  Length 117;
Best Local Similarity 60.0%  Pred. No. 9,50-26;
Matches 54: Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 52 TQPRFVNQHLGSHLVEALYLVCGERGFFYTPKT ----- 63
DB 28 TAPFVNQHLGSHLVEALYLVCGERGFFYTPKTRKEAEELQVNVHLENGKGAHQPEL 67

QY 86 -----RGVEQCTSLVQLENYCN 107
DB 86 ALEGLKRGVEQCTSLVQLENYCN 117

RESULT 9
US-09-947-563-4
: Sequence 4, Application US/09947563
: Patent No. US20020156234A1
: GENERAL INFORMATION:
: APPLICANT: Rabiner, Franz-Josef
: APPLICANT: Keller, Reinhold
: TITLE OF INVENTION: Improved process for obtaining
: TITLE OF INVENTION: insulin precursors having correctly bonded cysteine bridges
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
: Dunner
: STREET: 1300 I Street, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20005-3315
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible

```

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/947,563
: FILING DATE: 07-Sep-2001
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 09/134,836
: FILING DATE: <Unknown>
: ATTORNEY/AGENT INFORMATION:
: NAME: Leslie McDowell
: REGISTRATION NUMBER: 34,872
: REFERENCE/DOCKET NUMBER: 02481,1600-00000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 408-4000
: TELEFAX: (202) 408-4400
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 96 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: ORIGINAL SOURCE:
: ORGANISM: Escherichia coli
: FEATURE:
: NAME/KEY: Protein
: LOCATION: 1..96
: SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-947-563-4

Query Match      47.2%  Score 277;  DB 10;  Length 96;
Best Local Similarity 59.3%  Pred. No. 9,8e-25;
Matches 54: Conservative 0; Mismatches 3; Indels 34; Gaps 1;

QY 51 GTGPRFVNQHLGSHLVEALYLVCGERGFFYTPKT----- 85
DB 6 GNSARFVNQHLGSHLVEALYLVCGERGFFYTPKTRKEAEELQVNVHLENGKGAHQPEL 65

QY 86 -----RGVEQCTSLVQLENYCN 107
DB 86 ALEGLKRGVEQCTSLVQLENYCN 96

RESULT 10
US-09-736-611-12
: Sequence 12, Application US/09736611
: Patent No. US2001002159A1
: GENERAL INFORMATION:
: APPLICANT: Kjeldsen, Thomas
: APPLICANT: Ladvigsen, Svend
: APPLICANT: Kaarsholm, Niels
: TITLE OF INVENTION: Method For Making Insulin Precursors and
: TITLE OF INVENTION: Insulin Precursor Analogs
: FILE REFERENCE: 6056,200-US
: CURRENT APPLICATION NUMBER: US/09/736,611
: CURRENT FILING DATE: 2000-12-14
: PRIOR APPLICATION NUMBER: 60/181,443
: PRIOR FILING DATE: 2000-02-10
: PRIOR APPLICATION NUMBER: 60/211,441
: PRIOR FILING DATE: 2000-06-13
: PRIOR APPLICATION NUMBER: PA 1999 01868
: PRIOR FILING DATE: 1999-12-29
: PRIOR APPLICATION NUMBER: PA 2000 00440
: PRIOR FILING DATE: 2000-03-17
: NUMBER OF SEQ ID NOS: 18
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 12
: LENGTH: 124
: TYPE: PRT
: ORGANISM: N-terminal extension
US-09-736-611-12

```

```

1 LENGTH: 124
2 TYPE: PRT
3 ORGANISM: Artificial Sequence
4 FEATURE:
5 OTHER INFORMATION: Synthetic
6 US-09-894-711-12
7
8 Query Match 46.9%; Score 275.5; DB 10; Length 124;
9 Best Local Similarity 90.9%; Pred. No. 2.le-25;
10 Matches 50; Conservative 2; Mismatches 2; Indels 1; Gaps 1;
11
12 54 PRFVNOHLAGSHVLEALYLVCGERGFFFTPKT-RGIVECCCTSTICSLYOLENYCN 107
13 1:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
14 70 PRFVNOHLAGSHVLEALYLVCGERGFFFTDKDKGIVECCCTSTICSLYOLENYCN 124
15 1:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
16
17 RESULT 13
18 US-09-746-611-10
19 Sequence 10; Application US/09736611
20 Patient No. US2C010023069A1
21 GENERAL INFORMATION:
22 APPLICANT: Kjelsgen, Thomas
23 APPLICANT: Ludvigsen, Svend
24 APPLICANT: Kaarshoim, Niels
25 TITLE OF INVENTION: Method For Making Insulin Precursors and
26 TITLE OF INVENTION: Insulin Precursor Analogs
27 FILE REFERENCE: 6058.200-US
28 CURRENT APPLICATION NUMBER: US/09/736,611
29 CURRENT FILING DATE: 2000-12-14
30 PRIOR APPLICATION NUMBER: 60/181,443
31 PRIOR FILING DATE: 2000-02-10
32 PRIOR APPLICATION NUMBER: 60/211,441
33 PRIOR FILING DATE: 2000-06-13
34 PRIOR APPLICATION NUMBER: PA 1999 01868
35 PRIOR FILING DATE: 1999-12-29
36 PRIOR APPLICATION NUMBER: PA 2000 00440
37 PRIOR FILING DATE: 2000-03-17
38 NUMBER OF SEQ ID NOS: 18
39 SOFTWARE: FastSeq for Windows Version 4.0
40 SEQ ID NO 10
41 SEQ ID: 125
42 TYPE: PRT
43 ORGANISM: N-terminal extension
44 US-09-736-611-10
45
46 Query Match 46.9%; Score 275.5; DB 9; Length 125;
47 Best Local Similarity 90.9%; Pred. No. 2.le-25;
48 Matches 50; Conservative 2; Mismatches 2; Indels 1; Gaps 1;
49
50 54 PRFVNOHLAGSHVLEALYLVCGERGFFFTPKT-RGIVECCCTSTICSLYOLENYCN 107
51 1:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
52 71 PRFVNOHLAGSHVLEALYLVCGERGFFFTDKDKGIVECCCTSTICSLYOLENYCN 125
53 1:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
54
55 RESULT 14
56 US-09-740-359-10
57 Sequence 10; Application US/09740359
58 Patient No. US2C010041787A1
59 GENERAL INFORMATION:
60 APPLICANT: Kjelsgen, Thomas Borglum
61 APPLICANT: Ludvigsen, Svend
62 APPLICANT: Kaarshoim, Niels
63 TITLE OF INVENTION: Method for making insulin precursors and
64 TITLE OF INVENTION: insulin precursor analogues having improved fermentation
65 TITLE OF INVENTION: yield in yeast
66 FILE REFERENCE: 6143.200-US
67 CURRENT APPLICATION NUMBER: US/09/740,359
68 CURRENT FILING DATE: 2000-12-19
69 PRIOR APPLICATION NUMBER: PA 2000 00443
70 PRIOR FILING DATE: 2000-03-17
71 PRIOR APPLICATION NUMBER: PA 1999 01869
72 PRIOR FILING DATE: 1999-12-29
73 PRIOR APPLICATION NUMBER: 60/211,081
74 PRIOR FILING DATE: 2000-06-13

```


Search completed: September 16, 2003, 12:52:26
Job time : 34.7237 secs


```

F:1-30,31-51/Product: insulin #status experimental <ACH>
F:31-51/Domain: insulin chain A #status experimental <ACH>
F:7-37,19-50,36-41/Disulfide bonds: #status predicted

Query Match      46.6%; Score 273.5; DB 1; Length 51;
Best Local Similarity 96.2%; Pred. No. 8,7e-22;
Matches 50; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY 56 FVNHQCGSHLVEALYLVCGERGFFYTPKTRGVFCQCTSGSLYLENYCN 107
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 FVNHQCGSHLVEALYLVCGERGFFYTPKA-GIVRCCTSGSLYLENYCN 51

RESULT 3
INMRF
insulin - finback whale (tentative sequence)
C:Species: Balaeoptera physalus (finback whale, common rostrua)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C:Accession: A91918
R:Hama, H.; Titani, K.; Sakaki, S.; Narita, K.
C:Biochem. 56, 285-293, 1964
A:Title: The amino acid sequence in fin-whale insulin.
A:Reference number: A91918
A:Accession: A91918
A:Molecule type: Protein
C:Keywords: hormone; pancreas
C:Superfamily: insulin
F:1-30/Domain: insulin chain B #status experimental <ACH>
F:1-30,31-51/Product: insulin #status experimental <ACH>
F:31-51/Domain: insulin chain A #status experimental <ACH>
F:7-37,19-50,36-41/Disulfide bonds: #status predicted

Query Match      46.6%; Score 273.5; DB 1; Length 51;
Best Local Similarity 96.2%; Pred. No. 8,7e-22;
Matches 50; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY 56 FVNHQCGSHLVEALYLVCGERGFFYTPKTRGVFCQCTSGSLYLENYCN 107
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 FVNHQCGSHLVEALYLVCGERGFFYTPKA-GIVRCCTSGSLYLENYCN 51

RESULT 4
INEL
insulin - elephant
C:Species: Elephantidae gen. sp. (elephant)
C:Date: 24-Apr-1984 #sequence_revision 30-Sep-1984 #text_change 16-Jul-1999
C:Accession: A01584
R:Smith, L.F.
A:J. Med. 40, 662-665, 1966
A:Title: Species variation in the amino acid sequence of insulin.
A:Reference number: A90029; M:ID:66160119; PMID:5949523
A:Accession: A01584
A:Molecule type: Protein
A:Residues: 1-30,31-51 <SM>
A:Note: the species of elephant is not given, but it is most probably the Indian elephant
C:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-30/Domain: insulin chain B #status experimental <ACH>
F:1-30,31-51/Product: insulin #status experimental <ACH>
F:31-51/Domain: insulin chain A #status experimental <ACH>
F:7-37,19-50,36-41/Disulfide bonds: #status predicted

Query Match      46.6%; Score 273.5; DB 1; Length 51;
Best Local Similarity 94.3%; Pred. No. 8,7e-22;
Matches 49; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 56 FVNHQCGSHLVEALYLVCGERGFFYTPKTRGVFCQCTSGSLYLENYCN 107
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 FVNHQCGSHLVEALYLVCGERGFFYTPKTR-GIVRCCTSGSLYLENYCN 51

RESULT 5

```

```

B42179
insulin precursor - green monkey
C:Species: Cercopithecus aethiops (green monkey, grivet)
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 16-Jul-1999
C:Accession: B42179; A05232; S16494; S22056
R:Seino, S.; Bell, G.I.; Li, W.H.
Mol. Biol. Evol. 9, 193-203, 1992
A:Title: Sequences of primate insulin genes support the hypothesis of a slower rate
A:Reference number: A42179; M:ID:9221953; PMID:1560757
A:Accession: B42179
A:Molecule type: DNA
A:Residues: 1-110 <SEI>
A:Cross-references: EMBL:X61092; NID:922808; PIDN:CAA43405.1; PID:q22809
A:Note: sequence extracted from NCBI backbone (NCBIN:95185, NCBIF:95194)
R:Peterson, J.D.; Nehrllich, S.; Oyer, P.E.; Steiner, D.F.
J. Biol. Chem. 247, 4866-4871, 1972
A:Title: Determination of the amino acid sequence of the monkey, sheep, and dog proi
A:Reference number: A92111; M:ID:72258026; PMID:4626369
A:Accession: A05232
A:Molecule type: protein
A:Residues: 57-87 <PET>
C:Genetics:
A:Introns: 63/1
C:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-54/Domain: insulin chain B #status predicted <BC>
F:55-89/Domain: insulin #status predicted <MAT>
F:57-87/Domain: connecting peptide #status experimental <CPEP>
F:90-110/Domain: insulin chain A #status predicted <ACH>
F:31-96,41-109,95-100/Disulfide bonds: #status predicted

Query Match      46.5%; Score 273; DB 2; Length 110;
Best Local Similarity 60.2%; Pred. No. 2,2e-21;
Matches 53; Conservative 0; Mismatches 1; Indels 34; Gaps 1;

QY 54 PFVNHQCGSHLVEALYLVCGERGFFYTPKTR-REAEQFQGVGVGGPGAGSLQPLAL 82
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 23 PAFVNHQCGSHLVEALYLVCGERGFFYTPKTR-REAEQFQGVGVGGPGAGSLQPLAL 82

QY 86 -----RGVSCCTSGSLYLENYCN 107
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 93 RSLQKRGVSCCTSGSLYLENYCN 110

RESULT 6
INMRF
insulin precursor - crab-eating macaque
C:Species: Macaca fascicularis (crab-eating macaque)
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 16-Jul-1999
C:Accession: JQ0178
R:Wettkam, W.; Groeber, J.; Leineweber, M.; Wengenmayer, F.; Winnacker, E.L.
Gene 157, 179-183, 1982
A:Title: The nucleotide sequence of cDNA coding for preproinsulin from the primate M
A:Reference number: JQ0178; M:ID:83080474; PMID:6184262
A:Accession: JQ0178
A:Molecule type: mRNA
A:Residues: 1-110 <WEI>
A:Cross-references: GB:J05336; NID:9342121; PIDN:AAA36645.1; PID:q342122
C:Superfamily: insulin
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-54,90-110/Product: insulin #status predicted <MAT>
F:55-89/Domain: insulin chain B #status predicted <BC>
F:55-89/Domain: insulin connecting C peptide #status predicted <CPT>
F:90-110/Domain: insulin chain A #status predicted <ACH>
F:31-96,43-109,95-100/Disulfide bonds: #status predicted

Query Match      46.5%; Score 273; DB 2; Length 110;
Best Local Similarity 60.2%; Pred. No. 2,2e-21;
Matches 53; Conservative 0; Mismatches 1; Indels 34; Gaps 1;

QY 54 PFVNHQCGSHLVEALYLVCGERGFFYTPKTR-REAEQFQGVGVGGPGAGSLQPLAL 82
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 23 PAFVQHLCGSHLVEALYVGCGERGEFFYTPKIRREARDTGVQGVLEGGHAGSLQULAL 82

QY 86 -----HGIVQVQCTSCSYLYQLENYCN 107
 IIII IIII IIII IIII IIII

Db 83 EGSQQRGIVQVQCTSCSYLYQLENYCN 110

RESULT 7

INHY
 Insulin - hamster
 C:Species: Cricetinae gen. sp. (hamster)
 C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
 C:Accession: A91456
 R:Neelon, F.A.; Delcher, H.K.; Steinman, R.; Isherwood, R.E.
 Fed. Proc. 32, 350, 1973
 A:Title: Structure of hamster insulin: comparison with a hamster insulin.
 A:Reference number: A91456
 A:Accession: A91456
 A:Molecule type: Protein
 A:Residues: 1-30,31-51 <NEE>
 C:Superfamily: insulin
 C:Keywords: hormone; pancreas
 F:1-30/Domain: insulin chain B #status experimental <NEE>
 F:1-30,31-51/Product: insulin #status experimental <MAT>
 F:31-51/Domain: insulin chain A #status experimental <ACH>
 F:7-37,19-50,36-41/Disulfide bonds: #status predicted

Query Match 45.3% Score 271.5; DB 1; Length 51;
 Best Local Similarity 92.3%; Pred. No. 3,7e-21;
 Matches 49; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 56 FVNHLCGSHLVEALYVGCGERGEFFYTPKRGIVQVQCTSCSYLYQLENYCN 107
 IIII IIII IIII IIII IIII IIII IIII IIII IIII

Db 1 FVNHLCGSHLVEALYVGCGERGEFFYTPKS-GIVVQVQCTSCSYLYQLENYCN 51

RESULT 8

INMSSP
 Insulin - Egyptian spiny mouse (tentative sequence)
 C:Species: Acromys caliginus (Egyptian spiny mouse)
 C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 31-Mar-2000
 C:Accession: A01591
 R:Bucelli, H.F.; Humbel, R.E.
 Hoppe-Seyler's Z. Physiol. Chem. 353, 443-450, 1972
 A:Title: Isolation and partial structural analysis of insulin from mouse (Mus musculus)
 A:Reference number: A01591; MUID:72189454; PMID:5626210
 A:Contents: composition
 A:Accession: A01591
 A:Molecule type: protein
 A:Residues: 1-30,31-51 <HJE>
 C:Superfamily: insulin
 C:Keywords: hormone; pancreas
 F:1-30/Domain: insulin chain B #status predicted <ACH>
 F:1-30,31-51/Product: insulin #status predicted <MAT>
 F:31-51/Domain: insulin chain A #status predicted <ACH>
 F:7-37,19-50,36-41/Disulfide bonds: #status predicted

Query Match 45.7% Score 268.5; DB 1; Length 51;
 Best Local Similarity 92.3%; Pred. No. 2,9e-21;
 Matches 48; Conservative 3; Mismatches 0; Indels 1; Gaps 1;

QY 56 FVNHLCGSHLVEALYVGCGERGEFFYTPKRGIVQVQCTSCSYLYQLENYCN 107
 IIII IIII IIII IIII IIII IIII IIII IIII IIII

Db 1 FVNHLCGSHLVEALYVGCGERGEFFYTPKS-GIVVQVQCTSCSYLYQLENYCN 51

RESULT 9

A59151
 Insulin precursor - jack bean (fragments)
 N:Alternate names: hypoglycemic agent; plant insulin
 C:Species: Canavalia ensiformis (jack bean)
 C:Date: 07-Dec-1999 #sequence_revision 07-Dec-1999 #text_change 16-Sep-1996
 C:Accession: B59151; A59151

R:Oliveira, A.E.A.; Macrado, O.L.T.; Gomes, V.M.; Xavier-Neto, J.; Pereira, A.C.P.;
 Protein Pept. Lett. 6, 15-21, 1999
 A:Title: Jack bean seed coat contains a protein with complete sequence homology to
 A:Reference number: A59151
 A:Accession: B59151
 A:Molecule type: protein
 A:Residues: 1-30 <MACB>
 A:Accession: A59151
 A:Molecule type: Protein
 A:Residues: 31-51 <MACA>
 C:Comment: The two chains are probably produced from the same precursor.
 C:Superfamily: insulin
 F:1-30,31-51/Product: insulin #status experimental <MAT>
 F:1-30/Domain: chain B #status experimental <NEE>
 F:1-51/Domain: chain A #status experimental <ACH>
 F:7-37,19-50,36-41/Disulfide bonds: #status predicted

Query Match 45.6% Score 267.5; DB 2; Length 51;
 Best Local Similarity 92.3%; Pred. No. 3,7e-21;
 Matches 48; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 56 FVNHLCGSHLVEALYVGCGERGEFFYTPKRGIVQVQCTSCSYLYQLENYCN 107
 IIII IIII IIII IIII IIII IIII IIII IIII IIII

Db 1 FVNHLCGSHLVEALYVGCGERGEFFYTPKA-GIVVQVQCTSCSYLYQLENYCN 51

RESULT 10

IPSE
 Insulin precursor [validated] - human
 N:Alternate names: preproinsulin
 C:Species: Homo sapiens (man)
 C:Date: 24-Oct-1981 #sequence_revision 23-Oct-1981 #text_change 08-Dec-2000
 C:Accession: A93222; A94253; A93216; A93144; A92075; A91186; I58114; A01579
 R:Reis, G.; Pictet, R.L.; Rutter, W.J.; Cordell, B.; Fischer, E.; Goodman, H.M.
 Nature 284, 26-32, 1980
 A:Title: Sequence of the human insulin gene.
 A:Reference number: A93222; MUID:80120725; PMID:6243748
 A:Accession: A93222
 A:Molecule type: DNA
 A:Residues: 1-110 <REL>
 A:Cross-references: GB:J00265; NID:q186429; PIDN:AAA59172.1; PID:g386828
 R:Ulrich, A.; Duell, I.J.; Gray, A.; Brosius, J.; Sures, I.
 Science 209, 612-615, 1980
 A:Title: Genetic variation in the human insulin gene.
 A:Reference number: A94253; MUID:80236313; PMID:6246962
 A:Accession: A94253
 A:Molecule type: DNA
 A:Residues: 1-110 <GUL>
 A:Cross-references: GB:J00265; NID:q186429; PIDN:AAA59172.1; PID:g386828
 R:Reis, G.; Swain, W.F.; Pictet, R.; Cordell, B.; Goodman, H.M.; Rutter, W.J.
 Nature 282, 525-527, 1979
 A:Title: Nucleotide sequence of a cDNA clone encoding human preproinsulin.
 A:Reference number: A93216; MUID:80054779; PMID:503234
 A:Accession: A93216
 A:Molecule type: mRNA
 A:Residues: 1-110 <REL>
 A:Cross-references: GB:J00265; NID:q186429; PIDN:AAA59172.1; PID:g386828
 R:Sures, I.; Goodell, D.V.; Gray, A.; Ulrich, A.
 Science 268, 57-59, 1980
 A:Title: Nucleotide sequence of human preproinsulin complementary DNA.
 A:Reference number: A94251; MUID:80147417; PMID:6927840
 A:Accession: A94251
 A:Molecule type: mRNA
 A:Residues: 1-110 <SUR>
 A:Cross-references: GB:J00265; NID:q186429; PIDN:AAA59172.1; PID:g386828
 R:Nicol, D.S.H.W.; Smith, I.F.
 Nature 187, 483-485, 1960
 A:Title: Amino-acid sequence of human insulin.
 A:Reference number: A93144
 A:Accession: A93144
 A:Molecule type: Protein
 A:Residues: 25-54;96-110 <NIC>
 R:Oyer, P.E.; Cho, S.; Peterson, J.D.; Steiner, D.F.

```
J. Biol. Chem. 246, 1375-1386, 1971
A:Title: Studies on human proinsulin: isolation and amino acid sequence of the human pan
A:Reference number: A92075; MUID:71116450; PMID:5501771
A:Accession: A92075
A:Molecule type: protein
A:Residues: 57-87 <OE>
R:Ko, A.; Smyth, D.G.; Markussen, J.; Sundby, F.
Eur. J. Biochem. 20, 190-199, 1971
A:Title: Amino acid sequence of the C-peptide of human proinsulin
A:Reference number: A91186; MUID:71257722; PMID:5560424
A:Accession: A91186
A:Molecule type: protein
A:Residues: 57-87 <OA>
R:Lucassen, A.M.; Jullier, C.; Horessi, J.P.; Roillard, G.; Frommel, P.; Lathrop, M.; Bell
Nature Genet. 4, 305-310, 1993
A:Title: Susceptibility to insulin dependent diabetes mellitus maps to a 4.1 kb segment
A:Reference number: I58114; MUID:93364428; PMID:8458440
A:Accession: I58114
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-59,63-110 <RFS>
A:Cross-references: GB:115440; NID:q307371; PIDN:AAA59179.1; PID:q307372
R:Sieber, P.; Kamber, B.; Hartmann, A.; Joehl, A.; Rinkler, B.; Rittell, W.
Helv. Chim. Acta 57, 2617-2621, 1974
A:Title: Totalsynthese von Humaninsulin unter gezielter Bildung der Disulfidbindungen
A:Reference number: A91636; MUID:75077277; PMID:4443293
A:Contents: annotation; synthesis
A:Note: disulfide-bonded human insulin was synthesized; the synthetic hormone was found
R:Naithani, V.K.
Hoppe-Seyler's Z. Physiol. Chem. 354, 659-672, 1973
A:Title: The synthesis of C-peptide of human proinsulin
A:Reference number: A91658; MUID:75040007; PMID:4404564
A:Contents: annotation; synthesis of residues 57-87
R:Geiger, R.; Jaeger, G.; Koenig, W.
Chem. Ber. 105, 2347-2352, 1973
A:Title: Synthesis of the complete sequence of human proinsulin C-peptide and its C-pep
A:Reference number: A90914
A:Contents: annotation; synthesis of residues 57-87
R:Kaufmann, J.E.; Iminger, J.C.; Halban, P.A.
Biochem. J. 310, 869-874, 1995
A:Title: Sequence requirements for proinsulin processing at the P-chain/C-peptide juncti
A:Reference number: S58661; MUID:96013185; PMID:7575420
A:Contents: annotation; site-directed mutagenesis study of proinsulin processing
C:Genetics:
A:Gene: GDR:INS
A:Cross-references: GDR:119343; OMIM:176730
A:Map position: 11p15.5 11p15.5
A:Mutations: 63/
C:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-24/Domain: signal sequence #status experimental <SCH>
F:25-54/Domain: insulin chain B #status experimental <SCH>
F:55-54,90-110/Product: insulin #status experimental <MAT>
F:57-87/Domain: connecting C-peptide #status experimental <CPPT>
F:90-110/Domain: insulin chain A #status experimental <ACH>
F:31-96,43-109,95-100/Disulfide bonds: #status experimental

Query Match 45.5% Score 267; DB 1; Length 110;
Best Local Similarity 60.5% Pred. No. 9,3e-21;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 56 FVNQHLGSHIVEALYLVCGERGFFYTPKT----- 85
DB 25 FVNQHLGSHIVEALYLVCGERGFFYTPKTRREADLQGVGVGGAGSLQPLALEG 84
QY 86 ----RGIVQCCCTSCISLYOLENYCN 107
DB 85 SLOKRGIVQCCCTSCISLYOLENYCN 110

RESULT 12
insulin - sei whale
C:Species: Balanoptera borealis (sei whale)
C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 16-Jul-1999
C:Accession: A01582
R:Shihara, Y.; Saito, T.; Ito, Y.; Fujino, M.
Nature 361, 1468-1469, 1998
A:Title: Structure of sperm- and sei-whale insulins and their breakdown by whale pe
A:Reference number: A93142
A:Accession: A01582
A:Molecule type: protein
A:Residues: 1-30;31-51 <ISH>
C:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-30/Domain: insulin chain B #status experimental <ACH>
F:31-30,31-51/Product: insulin #status experimental <MAT>
F:51-51/Domain: insulin chain A #status experimental <ACH>
F:37-37,19-50,35-41/Disulfide bonds: #status predicted

Query Match 44.9% Score 263.5; DB 1; Length 51;
Best Local Similarity 92.3% Pred. No. 9.6e-21;
Matches 48; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 56 FVNQHLGSHIVEALYLVCGERGFFYTPKTRGIVEGCTSCISLYOLENYCN 107
DB 1 FVNQHLGSHIVEALYLVCGERGFFYTPKA-GIVEGCCASTCSLYOLENYCN 51

RESULT 13
insulin - goat
C:Species: Capra aegagrus hircus (domestic goat)
C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 16-Jul-1999
C:Accession: A01586
R:Smith, L.P.
Am. J. Med. 40, 662-666, 1966
A:Title: Species variation in the amino acid sequence of insulin
A:Reference number: A90029; MUID:66160119; PMID:5949593
A:Accession: A01586
A:Molecule type: protein
A:Residues: 1-30;31-51 <SM1>
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model:

Run on: September 16, 2003, 12:33:30 : Search time 8.32/85 Seconds
(without alignments)
604,297 Million cell updates/sec

Title: US-09-423-100-6

Perfect score: 567

Sequence:

1 MFPIFLSLPUNPMLRAHP.....IVEGQGISLCELYLNVN 107

Scoring table: ELOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127853 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127860

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	273.5	46.6	51	INS_BALPH	P01312 balneopter
2	273.5	46.6	51	INS_ELEMA	P01316 elephas max
3	273	46.5	110	INS_CERAE	P30407 ceropithes
4	273	46.5	110	INS_MACFA	P30406 macaca fasc
5	268.5	45.7	51	INS_ACCOA	P01324 acorys cabi
6	267	45.5	110	INS_HUMAN	P01368 homo sapien
7	267	45.5	110	INS_PANTR	P30410 pan troglod
8	266	45.3	110	INS_SPEPR	P01313 spermophilu
9	263.5	44.9	51	INS_FALBO	P01314 balneopter
10	263.5	44.9	51	INS_CABOR	P01320 camelus dro
11	263.5	44.9	51	INS_CABHI	P01319 capra airo
12	263	44.8	108	INS_FIG	P01315 sus scrofa
13	263	44.8	110	INS_RABIT	P01311 corytolius
14	262.5	44.7	51	INS_FELCA	P06306 felis stico
15	262	44.6	110	INS_CANFA	P01321 canis fami
16	260	44.3	110	INS_CRILA	P01317 crinophila
17	258.5	44.0	105	INS_BOVIN	P01317 los taure
18	257	43.8	108	INS_AOTIE	P00604 gorilla tri
19	257	43.8	110	INS_PSAOB	P02587 psammomya o
20	256.5	43.7	51	INS_DIDMA	P08169 didelphys m
21	255.5	43.5	217	SOMA_HUMAN	P02541 homo sapien
22	255.5	43.5	217	SOMA_MACMO	P33333 macaca mul
23	255.5	43.5	217	SOMA_PANTR	P58756 pan troglod
24	254.5	43.4	105	INS_SREBP	P01318 ovis arie
25	253	42.9	86	INS_HORSE	P01310 equus cabi
26	251.5	42.8	51	INS_CHARR	P01327 chinchilla
27	250	42.6	108	INS_MOUSE	P01325 mus muscul
28	249	42.4	110	INS1_RAT	P01322 ratu
29	249	42.4	217	SOMA_CALJA	P08141 callositax
30	249	42.4	217	SOMA_SALB3	P58141 salicib
31	248.5	42.3	51	INS_ANSAN	P07454 anser anser
32	248	42.2	110	INS2_MOUSE	P01326 mus muscul
33	248	42.2	110	INS2_RAT	P01323 ratu

34 246 41.9 52 1 INS_ACIGU P01423 acipenser g
35 245 41.7 103 1 INS_SELRF P01463 selasphorus
36 244.5 41.7 51 1 INS_HYSCR P01328 hystrix cri
37 244.5 41.7 51 1 INS_TRASC P01887 trachemys s
38 239.5 40.8 107 1 INS_CHICK P01332 gallus gall
39 236 40.2 217 1 SOM2_PANTR P58757 pan troglod
40 235.5 39.1 51 1 INS_ORNAN P09597 ornithotyr
41 233.5 39.8 81 1 INS_ANAPL P01333 anas platyr
42 231.5 39.4 51 1 INS_ALUMI P02703 alligator m
43 231 39.4 52 1 INS_LEPSP P09476 lepisosteus
44 228.5 38.9 51 1 INS_ZAODH P12708 zaocys dhum
45 228 38.8 217 1 SOM2_HUMAN P01242 homo sapien

ALIGNMENTS

RESULT 1
INS_BALPH STANDARD: PRT: 51 AA.
AC P01312
CT 21-JUL-1986 (Rel. 31, Created)
CT 21-JUL-1986 (Rel. 31, Last sequence update)
CT 21-JUL-1986 (Rel. 34, Last annotation update)
DE Insulin.
GN INS.
OR Balneoptera physalus (Pinback whale) (Common roqual), and
OR Physoter catodon (Sperm whale) (Physoter macrocephalus).
OC Eukaryota: Metazoa: Chordata: Cranialia: Vertebrata: Euteleostomi:
OC Mammalia: Eutheria: Cetartiodactyla: Cetacea: Mysticeti:
OC Balneopteridae: Balneoptera.
OX NCBI_TaxID=9770, 9755;
ON
RP PARTIAL SEQUENCE.
RC SPECIES=B.physalus;
KA Rama R., Titan K., Sakaki S., Narita K.;
KT "The amino acid sequence in fin-whale insulin.";
EL J. Biochem. 56:285-293(1964).
EN (2).
RP SEQUENCE.
RC SPECIES=P.catodon;
RA Ishihara Y., Saito T., Ito Y., Fujino M.;
RT "Structure of sperm- and sei-whale insulins and their breakdown by
RT whale pepsin.";
RL Nature 161:1468-1469(1958).
RN (3).
RP SEQUENCE.
RC SPECIES=P.catodon;
RA Harris J.I., Sander P., Naughton M.A.;
RT "Species differences in insulin.";
RL Arch. Biochem. Biophys. 65:427-438(1956).
OC FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
OC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
OC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
OC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
OC SUBUNIT: HETEROFORMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
OC DISULFIDE BONDS.
OC 1- SUBCELLULAR LOCATION: Secreted.
OC 1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
PIR: A91918; INWHIP.
PIR: A91142; INWHIP.
DR HSSP: P01317; IAPH.
DR InterPro: IPR004325; Ins/IGF/relax.
DR SMART: SM00078; IIGF. 1
DR PROSITE: PS00262; INSULIN; 1.
KW Insulin family; Hormone; Glucose metabolism.
FT CHAIN 1 30 INSULIN B CHAIN.
FT NON_CONS 30 31
FT CHAIN 31 51 INSULIN A CHAIN.
FT DISULFID 7 37 INTERCHAIN.
FT DISULFID 13 50 INTERCHAIN.
FT DISULFID 36 41
SQ SEQUENCE 51 AA: 5766 MW: 9007514691A7CDD CRC64:

```

Query Match      46.6%; Score 273.5; DS 1; Length 110;
Best Local Similarity 96.2%; Pred. No. 166-22;
Matches 50; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY 56 FVNQHLGGSHLVLEALYLVCGGGEFFYTKRAGVVEQCTSTCSLYQLENYCN 107
DB 1 FVNQHLGGSHLVLEALYLVCGGGEFFYTKRAGVVEQCTSTCSLYQLENYCN 5;

RESULT 2
INS_ELEMA STANDARD: PRT: 51 AA.
AC P01316;
DI 21-JUL-1986 (Rel. 01, Created)
DI 01-FEB-1996 (Rel. 01, Last sequence update)
DI 01-FEB-1996 (Rel. 33, Last annotation update)
DE Insulin.
GN INS.
OS Elephas maximus (Indian elephant).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Proboscidea; Elephantidae; Elephas.
OX NCBI_TaxID=9783;
RN [1]
RP SEQUENCE.
RX MEDLINE=66160119; PubMed=5049593;
RA Smith L.F.;
RL Am. J. Med. 40:662-666(1966).
CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X6:092; CAA43405.1;
CC FTK: B42179; B42179.
CC HSP: P01308; IAI0.
CC InterPro: IPR004825; Ins/IGF/relax.
CC SMART: PF000649; Insulin; 1.
CC SMART: SMC0078; IIGF; 1.
CC PROSITE: PS00262; INSULIN; 1.
CC Insulin family; Hormone; Glucose metabolism; Signal.
CC CHAIN 1 30 INSULIN B CHAIN.
CC NON_CONS 30 31
CC CHAIN 31 51 INSULIN A CHAIN.
CC CHAIN 51 52 INTERCHAIN.
CC DISULFID 7 37
CC DISULFID 19 50 INTERCHAIN.
CC DISULFID 36 41
CC SEQUENCE 51 AA; 5752 MW; 900755008457190 CRC64;

Query Match      46.6%; Score 273.5; DS 1; Length 110;
Best Local Similarity 96.2%; Pred. No. 166-22;
Matches 45; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY 56 FVNQHLGGSHLVLEALYLVCGGGEFFYTKRAGVVEQCTSTCSLYQLENYCN 107
DB 1 FVNQHLGGSHLVLEALYLVCGGGEFFYTKRAGVVEQCTSTCSLYQLENYCN 5;

RESULT 3
INS_CERAE STANDARD: PRT: 110 AA.
AC P30407; P01309;
DI 01-APR-1993 (Rel. 25, Created)
DI 01-APR-1993 (Rel. 25, Last sequence update)
DI 01-FEB-1996 (Rel. 33, Last annotation update)
DE Insulin precursor.
GN INS.
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;

```

```

Cercopithecoidea; Cercopithecidae.
NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92219953; PubMed 1560757;
FA Seino S., Bell G.L., Li W.;
FT "sequences of primate insulin genes support the hypothesis of a
FT slower rate of molecular evolution in humans and apes than in
FT monkeys.";
FL Mol. Biol. Evol. 9:193-203(1992).
RN [2]
RP SEQUENCE OF 57-87.
RX MEDLINE=72258016; PubMed=4626369;
FA Peterson C.D., Nehrlich S., Czer P.E., Steiner D.F.;
FT "determination of the amino acid sequence of the monkey, sheep, and
FT dog proinsulin C-peptides by a semi-micro Edman degradation
FT procedure.";
FL J. Biol. Chem. 247:4866-4871(1972).
CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X6:092; CAA43405.1;
CC FTK: B42179; B42179.
CC HSP: P01308; IAI0.
CC InterPro: IPR004825; Ins/IGF/relax.
CC SMART: PF000649; Insulin; 1.
CC SMART: SMC0078; IIGF; 1.
CC PROSITE: PS00262; INSULIN; 1.
CC Insulin family; Hormone; Glucose metabolism; Signal.
CC CHAIN 1 24 INSULIN B CHAIN.
CC NON_CONS 25 54
CC CHAIN 54 57 C PEPTIDE.
CC CHAIN 57 87
CC CHAIN 90 110 INSULIN A CHAIN.
CC CHAIN 110 111 INTERCHAIN.
CC DISULFID 31 96
CC DISULFID 43 104 INTERCHAIN.
CC DISULFID 95 100
CC SEQUENCE 110 AA; 12319 MW; 95A1F54BE7B24749 CRC64;

Query Match      46.5%; Score 273; DS 1; Length 110;
Best Local Similarity 96.2%; Pred. No. 41e-22;
Matches 53; Conservative 0; Mismatches 1; Indels 34; Gaps 1;

QY 54 PRVNVHLCGSHLVLEALYLVCGGGEFFYTKRAGVVEQCTSTCSLYQLENYCN 107
DB 23 PRVNVHLCGSHLVLEALYLVCGGGEFFYTKRAGVVEQCTSTCSLYQLENYCN 82

QY 86 -----RGIVVEQCTSTCSLYQLENYCN 107
DB 83 EGSLOKRGIVVEQCTSTCSLYQLENYCN 110

RESULT 4
INS_MACFA STANDARD: PRT: 110 AA.
AC P30406; P01309;
DI 21-JUL-1986 (Rel. 01, Created)
DI 13-AUG-1987 (Rel. 05, Last sequence update)
DI 01-FEB-1996 (Rel. 33, Last annotation update)
DE Insulin precursor.

```


RN SEQUENCE FROM N.A.
 RP MEDLINE=82147417; PubMed=6927840;
 RA Sures I., Goeddel D.V., Gray A., Ullrich A.;
 RT "Nucleotide sequence of human proinsulin complementary DNA";
 RL Science 238:57-59(1989).
 RN [5].
 RP SEQUENCE FROM N.A.
 RP MEDLINE=93364428; PubMed=3586640;
 RA Lucassen A.M., Bell J.I., Joller G., Lathrop M.;
 RT "Susceptibility to insulin dependent diabetes mellitus maps to a 1.1
 kb segment of DNA spanning the insulin gene and associated VNTR";
 RL Nat. Genet. 4:305-310(1994).
 RN [6].
 RP SEQUENCE FROM N.A.
 RP TISSUE=Pancreas;
 RC MEDLINE=22389257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenman C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Euetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hooq B.,
 RA Stapleton M., Soares M.B., Donalio M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci R., Prance G.,
 RA Raha S.S., Loquellaro N.A., Peters G.J., Abranson P.D., Mullighy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gurnatane P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay S.J., Hulyx S.W.,
 RA Villalón D.K., Kuzny D.M., Sedegren E.J., Lu X., Gibbs K.A.,
 RA Fahy J., Helton A., Kettunen M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bonfield W.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywicki M.I., Skalska H., Smalius D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [7].
 RP SEQUENCE OF 1-59 FROM N.A.
 RP TISSUE=Blood;
 RA Fajardy I.L., Wells J.D., Stackens C.G., Panse J.M.P.;
 RT "Description of a novel krill diallelic polymorphism (-127 T>G) C>G)
 within the 5' region of insulin gene";
 RL Submitted (JUL-1998) to the ENR/GenBank/EMBL databases.
 RN [8].
 RP SEQUENCE OF 25-54 AND 90-110.
 RA Nicol D.S.H.W., Smith G.F.;
 RT "Amino-acid sequence of human insulin";
 RL Nature 187:483-485(1960).
 RN [9].
 RP SEQUENCE OF 57-87.
 RX MEDLINE=71115410; PubMed=5101771;
 RA Oyer P.E., Cho S., Peterson J.D., Steiner D.F.;
 RT "Studies on human proinsulin. Isolation and amino acid sequence of
 the human pancreatic C-peptide";
 RL J. Biol. Chem. 246:1375-138(1971).
 RN [10].
 RP SEQUENCE OF 57-87.
 RX MEDLINE=71257722; PubMed=5560404;
 RA Ko A., Smyth D.G., Markusson J., Sundby E.;
 RT "The amino acid sequence of the C-peptide of human proinsulin";
 RL Eur. J. Biochem. 20:190-193(1971).
 RN [11].
 RP SYNTHESIS.
 RX MEDLINE=75077277; PubMed=4443293;
 RA Sieber P., Kamber B., Hartmann A., Geor A., Beniker B., Kitterl W.;
 RT "Total synthesis of human insulin under directed formation of the
 disulfide bonds";
 RL Helv. Chim. Acta 57:2617-2621(1974).
 RN [12].
 RP SYNTHESIS OF 57-87.
 RX MEDLINE=75040007; PubMed=48(3504);
 RA Nathani V.K.;
 RT "Studies on polypeptides. IV. The synthesis of C-peptide of human
 proinsulin";
 RL Hoppe-Seyler's Z. Physiol. Chem. 354:659-672(1973).
 RN [13].
 RP SYNTHESIS OF 85-69 AND 70-73.
 RX MEDLINE=73161253; PubMed=4598555;
 RA Geiger R., Volk A.;
 RT "Synthesis of peptides with the properties of human proinsulin C
 peptides (hC peptide). 3. Synthesis of the sequences 14-17 and 9-13
 of human proinsulin C peptides";
 RL Chem. Ber. 106:199-205(1973).
 RN [14].
 RP SYNTHESIS OF 84-87.
 RX MEDLINE=73161261; PubMed=4598553;
 RA Geiger R., Jaeger G., Keonig W., Treuth G.;
 RT "Synthesis of peptides with the properties of human proinsulin C
 peptides (hC peptide). 1. Scheme for the synthesis and preparation of
 the sequence 28-31 of human proinsulin C peptide";
 RL Chem. Ber. 106:188-192(1973).
 RN [15].
 RP VARIANT LOS ANGELES SER-48.
 RX MEDLINE=84016053; PubMed=6312455;
 RA Haedea M., Chan S.J., Kwok S.C.M., Rubenstein A.H., Steiner D.F.;
 RT "Studies on mutant human insulin genes: identification and sequence
 analysis of a gene encoding [SerB24]insulin";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:6366-6370(1983).
 RN [16].
 RP VARIANTS LOS ANGELES SER-48 AND CHICAGO LEU-49.
 RX MEDLINE=84170233; PubMed=6424111;
 RA Shoelson S., Fickova M., Haneda M., Nahum A., Musso G., Kaiser E.L.,
 RA Rubenstein A.H., Tager H.;
 RT "Identification of a mutant human insulin predicted to contain a
 serine-for-phenylalanine substitution";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:7390-7394(1983).
 RN [17].
 RP VARIANT PROVIDENCE ASP-34.
 RX MEDLINE=87175640; PubMed=3470784;
 RA Chan S.J., Seino S., Gruppiso P.A., Schwartz K., Steiner D.F.;
 RT "A mutation in the B chain coding region is associated with impaired
 proinsulin conversion in a family with hyperproinsulinemia";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:2194-2197(1987).
 RN [18].
 RP VARIANT WAKAYAMA LEU 92.
 RX MEDLINE=87036122; PubMed=3537011;
 RA Sakura H., Iwamoto Y., Sakamoto Y., Kuzuya T., Hirata H.;
 RT "Structurally abnormal insulin in a diabetic patient. Characterization
 of the mutant insulin A3 (Val-->Leu) isolated from the pancreas";
 RL J. Clin. Invest. 78:1666-1672(1986).
 RN [19].
 RP VARIANT HIS-89.
 RX MEDLINE=90317021; PubMed=2196279;
 RA Barbelli F., Raben N., Kadowaki T., Cama A., Accili D., Gabbay K.H.,
 RA Merenich J.A., Taylor S.L., Roth J.;
 RT "Two unrelated patients with familial hyperproinsulinemia due to a
 mutation substituting histidine for arginine at position 65 in the
 proinsulin molecule: identification of the mutation by direct
 sequencing of genomic deoxyribonucleic acid amplified by polymerase
 chain reaction";
 RL J. Clin. Endocrinol. Metab. 71:164-169(1990).
 RN [20].
 RP VARIANT HIS-89.
 RX MEDLINE=85261946; PubMed=4019786;
 RA Shibasaki Y., Kawakami T., Kanazawa Y., Akanuma Y., Takaku F.;
 RT "Posttranslational cleavage of proinsulin is blocked by a point
 mutation in familial hyperproinsulinemia";
 RL J. Clin. Invest. 76:378-380(1985).
 RN [21].
 RP VARIANT KYOTO LEU-89.
 RX MEDLINE=92291307; PubMed=1601997;
 RA Yano H., Kitano N., Morimoto M., Polonsky K.S., Izura H., Seino Y.;
 RT "A novel point mutation in the human insulin gene giving rise to
 hyperproinsulinemia (proinsulin Kyoto)";
 RL J. Clin. Invest. 89:1902-1907(1992).
 RL


```

Pfam: PF00049; Insulin; 1
PRINTS: P800277; INSULIN;
SMART: SM00028; IIGF; 1
PROSITE: PS00262; INSULIN; 1
KW Insulin family; Hormone; Glucose metabolism; Signal.
FT SIGNAL 1 24 BY SIMILARITY
FT CHAIN 25 54 INSULIN H CHAIN
FT PROPEP 57 87 C PEPTIDE
FT CHAIN 90 110 INSULIN A CHAIN
FT DISULFID 31 96 INTERCHAIN (BY SIMILARITY)
FT DISULFID 43 109 INTERCHAIN (BY SIMILARITY)
FT DISULFID 95 100 BY SIMILARITY
SQ SEQUENCE 110 AA; 12064 MW; 4511758066228EE5 CRC64;

Query Match 45.3%; Score 266; DB 1; Length 110;
Best Local Similarity 57.4%; Pred. No. 2.2e-21;
Matches 54; Conservative 1; Mismatches 3; Indels 36; Gaps 2;

QY 50 LGTGP--REVNHLCGSHLVEALYLVGCGRGGFFVTPKTRIGIVEOCGCTSCISLYOLENYCN 107
DB 17 LGPDPAQAFVNHLCGSHLVEALYLVGCGRGGFFVTPKTRIGIVEOCGCGOVELAAGPGASL 76
QY 86 -----RIVFOCCITSCISLYOLENYCN 107
DB 77 PQPLALEMALQKRGIVFOCCITSCISLYOLENYCN 110

RESULT: 9
INS_BALBO STANDARD: PRT: 51 AA.
AC P01314;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 01-OCT-1996 (Rel. 34, Last annotation update)
GN Insulin.
OS Balaenoptera borealis (Sei whale).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;
OC Balaenopteridae; Balaenoptera.
EX NCBI_TaxID=9768;
RN [1]
RP SEQUENCE.
RA Ishihara Y., Saito T., Ito Y., Fujino M.;
RL "Structure of sperm and sei whale insulins and their breakdown by
RT whale pepsin."
CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR PIR: A01582; INWEL.
DR HSSP: P01317; 1A9H.
DR InterPro: IPR004825; Ins/IGF/relax.
DR SMART: SM00078; IIGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
KW Insulin family; Hormone; Glucose metabolism.
FT CHAIN 1 30 INSULIN B CHAIN
FT NON_CONS 30 31 INSULIN A CHAIN
FT CHAIN 31 51 INTERCHAIN
FT DISULFID 7 37 INTERCHAIN
FT DISULFID 19 50 INTERCHAIN
FT DISULFID 36 41
SQ SEQUENCE 51 AA; 5693 MW; 901E88BA085A7DDD CRC64;

Query Match 44.9%; Score 263.5; DB 1; Length 51;
Best Local Similarity 90.4%; Pred. No. 1.9e-21;
Matches 47; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 56 FVNHLCGSHLVEALYLVGCGRGGFFVTPKTRIGIVEOCGCTSCISLYOLENYCN 107
DE 1 FANHLCGSHLVEALYLVGCGRGGFFVTPKRA-GIVEOCGASVSLYOLENYCN 51

RESULT: 11
INS_CAPHI STANDARD: PRT: 51 AA.
AC P01319;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 01-FEB-1996 (Rel. 33, Last annotation update)
GN Insulin.
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
EX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE.
RA Smith L.F.;

```

```

RT "Species variation in the amino acid sequence of insulin."
RL Am. J. Med. 40:662-666(1966).
CC -1- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -2- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -3- SUBCELLULAR LOCATION: Secreted.
CC -4- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR PIR: A01586; INGI.
DR HSSP: P01317; IAPB.
DR InterPro: IPR004825; Ins/IGF/relax.
DR SMART: SM00078; IIGF; 1.
DR PROSITE: PS00052; INSULIN; 1.
KW Insulin family; Hormone; Glucose metabolism.
FT CHAIN 1 30 INSULIN B CHAIN.
FT NON_CONS 40 31
FT CHAIN 31 51 INSULIN A CHAIN.
FT DISULFID 7 37 INTERCHAIN.
FT DISULFID 19 50 INTERCHAIN.
FT DISULFID 36 41
SQ SEQUENCE 51 AA; 5592 MW; 5037B50CDB4F720D CR454;

Query Match. 44.9%, Score 253.5; 56 21; 30000000
Best Local Similarity 90.4%, Pred. No. 1,90-217
Matches 47: Conservative 1; Mismatches 3; Indels 1; Gaps 1;

CY 56 FVNHLGGSHLEALNLCVCGRRFFFTYTKTGGIVEGCTSTSLNYGLENYCN 107
D6 111 111 111 111 111 111 111 111 111 111 111 111 111 111
1 FVNHLGGSHLEALNLCVCGRRFFFTYTKA GIVEGCTAGVSLYGLNENYCN 91

RESULT 12
ID INS_PIG STANDARD: PRT: 106 AA.
AC P01315; GTSU55;
DT 21-JUL-1986 (Rel. 01, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Insulin precursor.
GN INS.
OS Sus. scrofa (Pig).
CC Eukaryota; Metazoa; Chordata; Granata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
CX NCBI_TaxID:9823;
RN 1;
RC SEQUENCE FROM N.A.
RA Han X.G., Tsch B.E.;
RL "Complete porcine preproinsulin cDNA sequence."
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN 2;
RC SEQUENCE FROM N.A.
RA STRAIN: Large white;
RX MEDLINE=22135956; PubMed=12140686;
RA Amarger V., Nguyen M., Lacroix A.S., Braunschweig M., Nezer C.,
RA Georges M., Andersson L.;
RT "Comparative sequence analysis of the INS IGF2-IR4 gene cluster in
RT pigs."
RL Mann. Genome 13:388-398(2002).
RN 3;
RC SEQUENCE OF 25-108.
RX MEDLINE=6826485; PubMed 5657063;
RA Chance R.E., Ellis R.M., Broder W.W.;
RT "Porcine preproinsulin: characterization and amino acid sequence."
RL Science 161:165-167(1968).
RN 4;
RC REVISION TO 59.
RA Chance R.E.;
RL Submitted (JUL-1970) to the PIR data bank.
RN 5;
RX X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RA Blundell T.L., Dodson G.G., Hodgkin D., Moras D.;

```

```

RT "Insulin. The structure in the crystal and its reflection in
RT chemistry and biology."
RL Adv. Protein Chem. 26:279-402(1972).
RN 6;
RX X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
RA Isaacs N.W., Agarwal R.C.;
RT "Experience with fast Fourier least squares in the refinement of the
RT crystal structure of rhombohedral 2-zinc insulin at 1.5-A
RT resolution."
RL Acta Crystallogr. A 34:792-791(1978).
RN 7;
RX X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
RL MEDLINE=69099318; PubMed=2405485;
RA Baker E.N., Blundell T.L., Cutfield J.F., Cutfield S.M., Dodson E.J.,
RA Dodson G.G., Crowfoot Hodgkin D.M., Hubbard R.E., Isaacs N.W.,
RA Reynolds C.D., Sakabe K., Sakabe N., Vijayan N.M.;
RT "The structure of 2Zn pig insulin crystals at 1.5-A resolution."
RL Philos. Trans. R. Soc. Lond., B, Biol. Sci. 319:369-456(1989).
RN 8;
RX X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RL MEDLINE=92126260; PubMed=1772633;
RA Palschmidt P., Hansen F.B., Dodson E., Dodson G., Korber F.;
RT "Structure of porcine insulin cocrystallized with clupeine 2."
RL Acta Crystallogr. B 47:975-986(1991).
RN 9;
RX X-RAY CRYSTALLOGRAPHY.
RL MEDLINE=9122450; PubMed=2625410;
RA Badier J., Harris M.R., Reynolds C.D., Evans A.C., Dodson E.J.,
RA Dodson G.G., North A.C.T.;
RT "Structure of the pig insulin dimer in the cubic crystal."
RL Acta Crystallogr. B 47:127-136(1991).
RN 10;
RX X-RAY CRYSTALLOGRAPHY (1.65 ANGSTROMS).
RA Diao J.-S., Wan Z.-L., Chang W.-R., Liang D.-C.;
RT "Structure of monomeric porcine DesB1-B2 despitapeptide (B26-B30)
RT insulin at 1.65-A resolution."
RL Acta Crystallogr. D 53:507-512(1997).
CC -1- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -2- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -3- SUBCELLULAR LOCATION: Secreted.
CC -4- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC DATABASE: NAME-Protein Spotlight;
CC NOTE: Issue 9 of April 2001;
CC WWW: http://www.expasy.org/spotlight/articles/spt1009.html.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL: AF064555; AAC77920.1; ALT_INIT.
DR EMBL: AY044828; AAL69550.1;
DR PDB: 3INS; 09-JAN-89.
DR PDB: 4INS; 31-JUL-94.
DR PDB: 5INS; 31-JAN-94.
DR PDB: 7INS; 31-JAN-94.
DR PDB: 9INS; 15-OCT-91.
DR PDB: 112A; 15-OCT-91.
DR PDB: 112B; 15-OCT-91.
DR PDB: 27C1; 29-JAN-96.
DR PDB: 1MR1; 29-JAN-96.
DR PDB: 3MTS; 29-JAN-96.
DR PDB: 1DE1; 16-JUN-97.
DR PDB: 1SD6; 01-APR-98.
DR PDB: 1WAV; 28-FEB-99.
DR PDB: 12E1; 16-FEB-99.

```

```

DR PDB: 1ZM1: 28-JAN-98.
DR PDB: 1ZM1: 28-JAN-98.
DR InterPro: IPR004825: Ins/IGF/relax.
DR Pfam: PF00649: Insulin; 1.
DR SMART: SM00078: IIGF; 1.
DR PROSITE: PS00262: INSULIN; 1.
KW Insulin family; Hormone; Glucose metabolism; Secreted; 3D structure.
FT SIGNAL 1 24
FT CHAIN 25 54 INSULIN B CHAIN.
FT PROPEP 57 87 C PEPTIDE.
FT CHAIN 88 108 INSULIN A CHAIN.
FT DISULFID 91 94 INTERCHAIN.
FT DISULFID 43 107 INTERCHAIN.
FT DISULFID 93 98
FT HELIX 24 44
FT STRAND 48 48
FT STRAND 59 94
FT HELIX 100 106
FT STRAND 107 107
SQ SEQUENCE 108 AA: CB4591B429E58EBE CRC64:

Query Match 44.8%; Score 263; DB 1: Length 108;
Best Local Similarity 60.7%; Pred. No. 4.6e-21;
Matches 51; Conservative 0; Mismatches 1; Indels 32; Gaps 1.

QY 56 FVNHLCGSHVEALYLVCGRGFFYTPKT----- 85
DB 25 FVNHLCGSHVEALYLVCGRGFFYTPKARAEAFQAGAEVHGGJGGLOAALRGFF 84

QY 86 ---RGIVECCCTICSLYQLENYCN 107
DB 85 QRGIVECCCTICSLYQLENYCN 108

RESULT 13
INS_RABIT
ID INS_RABIT STANDARD: PRG: 110 AA.
AC F01311;
DT 2-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Insulin precursor.
GN INS.
CS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white; TISSUE=Pancreas;
RX MEDLINE=94179210; PubMed=8132571;
RA Devaskar S.J., Giddings S.J., Rajakumar P.A., Carnaqui L.R.,
RA Menon R.K., Zahm D.S.;
RT "Insulin gene expression and insulin synthesis in mammalian neuronal
RT cells.";
RJ J. Biol. Chem. 269:8445-8454(1994);
RN [2]
RP SEQUENCE OF 25-54 AND 90-110.
RX MEDLINE=66160119; PubMed=5945594;
RA Smith L.P.;
RT "Species variation in the amino acid sequence of insulin.";
RJ Am. J. Med. 40:662-666(1966).
RN [3]
RP SEQUENCE OF 56-110 FROM N.A.
RA Giddings S.J., Carnaqui L.R., Devaskar S.J.;
RP Submitted (APR-1991) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -!- SUBCELLULAR LOCATION: Secreted.

```

```

-!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
-!- This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement. (See http://www.isb-sib.ch/announce/
or send an email to licenc@sib-sib.ch).
EMBL: U03610; AAAL9C33.1;
EMBL: M01153; AAAL7546.1;
PIR: A33438; INRR.
SSP: P01308; IYIM.
InterPro: IPR004825; Ins/IGF/relax.
Pfam: PF00649; Insulin; 1.
SMART: SM00078; IIGF; 1.
PROSITE: PS00262; INSULIN; 1.
KW Insulin family; Hormone; Glucose metabolism; Signal.
FT SIGNAL 1 24
FT CHAIN 25 54 INSULIN B CHAIN.
FT PROPEP 57 87 C PEPTIDE.
FT CHAIN 90 110 INSULIN A CHAIN.
FT DISULFID 43 109 INTERCHAIN.
FT DISULFID 95 100
FT CONFLICT 83 83 F -> Y (JN REF. 3).
SQ SEQUENCE 110 AA: 82D2975985D77F8 CRC64:

Query Match 44.8%; Score 263; DB 1: Length 110;
Best Local Similarity 59.3%; Pred. No. 4.6e-21;
Matches 51; Conservative 1; Mismatches 0; Indels 34; Gaps 1.

QY 56 FVNHLCGSHVEALYLVCGRGFFYTPKT----- 85
DB 25 FVNHLCGSHVEALYLVCGRGFFYTPKSREVELOVQAGELGGGAGGLGPSALEL 84

QY 86 ---RGIVECCCTICSLYQLENYCN 107
DB 85 ALOKRGIVECCCTICSLYQLENYCN 110

RESULT 14
INS_FELCA
ID INS_FELCA STANDARD: PRG: 51 AA.
AC P06306;
DT 01-JAN-1986 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Insulin.
GN INS.
CS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE.
RX MEDLINE=86214076; PubMed=3518635;
RA Hadden G., Gavell G., Mutt V., Journvall H.;
RT "Characterization of cat insulin.";
RJ Arch. Biochem. Biophys. 247:20-27(1986).
CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
PIR: A01588; INCT.
HSSP: P01317; IAPH.
InterPro: IPR004825; Ins/IGF/relax.
PRINTS: PR00277; INSULINB.

```


Result No.	Query		Length	DB ID	Description
	Score	Match			
1	267	45.5	110	6	Q8XV2
2	266	45.3	110	1	Q91X73
3	251	42.8	110	6	Q9XN66
4	249	42.4	217	6	Q9XN60
5	228	38.8	245	4	Q14644
6	219.5	37.4	106	13	Q91XQ7
7	213	36.3	212	5	Q07468
8	213	36.5	217	5	Q07367
9	201.5	34.3	110	13	Q947A6
10	201	34.2	217	6	Q07469
11	197	33.6	217	4	Q14407
12	195.5	33.3	108	13	Q90CE5
13	195.5	33.3	108	13	Q90ZM4
14	195	33.2	111	13	Q90TB0
15	195	33.2	217	6	Q9XN69
16	193.5	33.0	110	13	Q90ZY1


```

DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Insulin.
OS Spermophilus tridecemlineatus (Thirteen-lined ground squirrel)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurimorphia; Sciuridae; Spermophiles
OC Spermophilus
OX NCBI_TaxID=44179;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RA Tredrea M.M., Buck M.J., Guhaniyogi J., Squire T.L., Andrews M.T.;
RT "Regulation of PK4 expression in a hibernating mammal.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SURCELLULAR LOCATION: SECRETED (by SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AY038604; AAK72556.1; -.
DR HSSP: P01306; ILNF.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin; 1.
DR PRINTS: PR00277; INSULIN.
DR SMART: SM00078; ILGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
SQ SEQUENCE 110 AA: 12064 MW: 45176806222885 SMC64;
Query Match 45.3%; Score 246; DB 6; Length 110;
Best Local Similarity 57.4%; Pred. No. 8,70-24;
Matches 54; Conservative 1; Mismatches 3; Indels 16; Gaps 2;

QY 50 LGTPG--RVVNHLCGSHVLYVCGERGFFYTKT----- 85
DB 17 LGPPAQAFVNVHLCGSHVLYVCGERGFFYTKSRREVVEQQGGQVFLGSGPGAGL 76
QY 86 -----RGIVEQCCTISGLYOLENYCN 107
DB 77 PPTALEMALKRGIVEQCCTISGLYOLENYCN 110

RESULT 3
Q8WNW6 PRELIMINARY: PRI: 110 AA.
AC Q8WNW6;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Preproinsulin.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felineae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RA Okamoto S., Morimatsu M.;
RT "cat insulin.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SURCELLULAR LOCATION: SECRETED (by SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AB043535; BAB04116.1; -.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin; 1.
DR PRINTS: PR00277; INSULIN.
DR SMART: SM00078; ILGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
SQ SEQUENCE 110 AA: 12069 MW: 95FB62170C7B52A1 SMC64;
Query Match 42.8%; Score 251; DB 6; Length 110;
Best Local Similarity 55.8%; Pred. No. 5,90-22;
Matches 48; Conservative 2; Mismatches 2; Indels 34; Gaps 1;

QY 56 FVNHLCGSHVLYVCGERGFFYTKT----- 85
DB 25 FVNHLCGSHVLYVCGERGFFYTKARREEDLQKIALGEPAGGLQPSALEA 84

```

```

QY 86 -----RGIVEQCCTISGLYOLENYCN 107
DB 85 PLOKRGVVECCASVCSLYOLENYCN 110

RESULT 4
Q8WNE0 PRELIMINARY: PRI: 217 AA.
AC Q8WNE0;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Growth hormone.
OS Gn-N.
OS Ateles geoffroyi (Black-handed spider monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Ateles.
OX NCBI_TaxID=9509;
RN [1]
RP SEQUENCE FROM N.A.
RA Revol A., Esquivel D., Santiago D., Barrera-Saldana H.;
RT "Independent duplication of the growth hormone gene in three Anthropoidae lineages.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF374234; AAL72286.1; -.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone; 1.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
SQ SEQUENCE 217 AA: 24894 MW: 4258299F41EBAAR6 CRC64;
Query Match 42.4%; Score 249; DB 6; Length 217;
Best Local Similarity 97.9%; Pred. No. 20-21;
Matches 47; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FPTIPLSRLFNAMLRHRLHQIAFDYQFEFAYIPKQKYSFLQNP 49
DB 27 FPTIPLSRLFNAMLRHRLHQIAFDYQFEFAYIPKQKYSFLQNP 74

RESULT 5
Q14644 PRELIMINARY: PRI: 245 AA.
AC Q14644;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Placental growth hormone isoform hGH-V3 precursor.
OS HCR-V.
OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Term placenta;
MEDLINE: 9637377; PubMed: 9709563;
RA Bouszewski C.D., Svensson B.A., Jansson T., Clark R.;
RA Carlsson L.M.S., Carlsson B.;
RL "Cloning of two novel growth hormone transcripts expressed in human placenta.";
DR J. Clin. Endocrinol. Metab. 83:2878-2885(1998).
DR EMBL: AF006061; AAB71829.1; -.
DR HSSP: P0241; IAZ2.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone; 1.
DR PROSITE: PS00265; SOMATOTROPIN_1; 1.
DR SIGNAL.
ET SIGNAL.
SQ SEQUENCE 245 AA: 27110 MW: 14CC7F8CD75D91C8 CRC64;
Query Match 38.8%; Score 228; DB 4; Length 245;

```



```

RT *Molecular cloning of preproinsulin cDNAs from several
RT Osteo-sarcomas and a chondroid.
RL Mol. Cell. Endocrinol. 174:51-58(2001).
CC -:- SUPCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -:- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: AF159586; AAK2972.1; -.
DR HSSP: P01408; IHS5.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00649; Insulin_1.
DR SMART: SM00074; IIGF_1.
DR PROSITE: PS00262; INSULIN.
SQ SEQUENCE 110 AA: 1242; MW: 36000.6590872ED6 CRC64:
Query Match 34.3%; Score 201.5; DB 1; Length 116;
Best Local Similarity 43.5%; Pred. No. 4e-16;
Matches 37; Conservative 8; Mismatches 5; Indels 15; Gaps 1;
QY 58 NCHLCGSHVYKALVYKSGRFFVTPK-----
DB 26 SCHLCGSHVLDALVWCKGKGFYCPKTKRVQVLDLFLSPKSAQENAEYFVFKQQL 43
QY 86 ---RIVROCCSTSCSLVQENYCN 107
DB 86 KVKRGIVEQCCHPCNIFQLQNYCN 110
RESULT 10
Q07369 PRELIMINARY; PRT: 217 AA.
AC Q07369;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DI 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DI 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Chorionic somatomammotropin-3.
OS Macaca mulatta (Phebus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE:Midpregnancy placenta;
FX MEDLINE-94008724; PubMed=8404617;
RA Golos T.G., Durning M., Fisher J.K., Fowler P.D.;
RT "Cloning of four growth hormone/chorionic somatomammotropin-related
RT complementary deoxyribonucleic acids differentially expressed during
RT pregnancy in the rhesus monkey placenta."
RL Endocrinology 133:1744-1752(1993).
DR EMBL: L16554; AAA18841.1;
DR HSSP: P01241; IAXI.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone_1;
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
SQ SEQUENCE 217 AA: 24874; MW: 27686.0704815185 CRC64:
Query Match 34.28; Score 261.1; DB 6; Length 217;
Best Local Similarity 74.53; Pred. No. 1e-15;
Matches 35; Conservative 8; Mismatches 4; Indels 7; Gaps 1;
QY 3 PTFVSLRFLDQNAHLRHLQAFDTYQEEFVYIPKQKYSPLQN 49
DB 28 PTVLSRLFDQNAHLRHLQAFDTYQEEFVYIPKQKYSPLQN 74
RESULT 11
Q14407 PRELIMINARY; PRT: 217 AA.
AC Q14407;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DI 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DI 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Chorionic somatomammotropin CS-2 (Chorionic somatomammotropin hormone-

```

```

DE Homo sapiens (human).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE-89307277; PubMed=2744760;
FX Chen K.Y., Iiao Y.C., Smith D.H., Herrera-Saldana H.A., Gellinas R.E.,
FX Seeburg P.H.;
RT "The human growth hormone locus: nucleotide sequence, biology, and
RT evolution."
RL Genomics 4:479-497(1989).
QY SEQUENCE FROM N.A.
DB MEDLINE-91102558; PubMed=1980158;
FX Vencak-Tones C.L., Phillips J.A. III.;
RT "Hot spots for growth hormone gene deletions in homologous regions
RT outside of Alu repeats."
RL Science 250:1745-1748(1990).
QY SEQUENCE FROM N.A.
DB TISSUE=Placenta;
FX Strausberg R.;
RT Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
FX Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: J03071; AAA52553.1; -.
DR EMBL: BC022044; AAH22044.1; -.
DR EMBL: BC035965; AAH35965.1; -.
DR HSSP: P01241; IAX2.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone_1;
DR PROSITE: PS00836; SOMATOTROPIN.
DR PROSITE: PS00286; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
SQ SEQUENCE 217 AA: 24994; MW: 39FAACDDB6B2E951 CRC64:
Query Match 43.6%; Score 197; DB 4; Length 217;
Best Local Similarity 80.8%; Pred. No. 3e-15;
Matches 36; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 4 TPTSLRFLDQNAHLRHLQAFDTYQEEFVYIPKQKYSPLQN 48
DB 29 TPTSLRFLDQNAHLRHLQAFDTYQEEFVYIPKQKYSPLQN 73
RESULT 12
Q14405 PRELIMINARY; PRT: 108 AA.
AC Q14405;
DT 01-MAR-2003 (TrEMBLrel. 16, Created)
DI 01-MAR-2003 (TrEMBLrel. 16, Last sequence update)
DI 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE zebu-in precursor.
RN [1]
RP Brachydanio rerio (Zebrafish) (Danio rerio).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE-99425150; PubMed=10495291;
FX Argenton F., Zecchin E., Bortolussi M.;
RT "Early appearance of pancreatic hormone-expressing cells in the
RT zebrafish embryo."
RL Mech. Dev. 87:217-221(1999).
QY -:- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -:- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.

```

2 FPTPLSRFLFONAMLRRAHRLHQLAFDITYQEFEAYIPKEQKYSFLQNP 49

Db 27 FBI/PLS/LEADS/HQ/HQVAND/VE/CELEBR/CI/PAK/KY/IN/ENP 24

Search completed: September 16, 2003, 12:40:00
Job time : 27.8132 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Computer Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 12:31:20 / Search time 44.358 seconds
(without alignments)
536.746 Million cell updates/sec

Title: US-09-423-100-7
Perfect score: 797
Sequence: 1 MFPTFLSRIFUNAMLAHR.....IVEGQSTSTCSVQLNVCN L50

Scoring table: BLOSUM62
Gap 10.0, Gapext 0.5

Searched: 1107863 seqs, 119726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match (%):
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

1:	/SIDS1/qcdata/qcseq/qcseq-emb1/AA1981.DAT:
2:	/SIDS1/qcdata/qcseq/qcseq-emb1/AA1982.DAT:
3:	/SIDS1/qcdata/qcseq/qcseq-emb1/AA1983.DAT:
4:	/SIDS1/qcdata/qcseq/qcseq-emb1/AA1984.DAT:
5:	/SIDS1/qcdata/qcseq/qcseq-emb1/AA1985.DAT:
6:	/SIDS1/qcdata/qcseq/qcseq-emb1/AA1986.DAT:
7:	/SIDS1/qcdata/qcseq/qcseq-emb1/AA1987.DAT:
8:	/SIDS1/qcdata/qcseq/qcseq-emb1/AA1988.DAT:
9:	/SIDS1/qcdata/qcseq/qcseq-emb1/AA1989.DAT:
10:	/SIDS1/qcdata/qcseq/qcseq-emb1/AA1990.DAT:
11:	/SIDS1/qcdata/qcseq/qcseq-emb1/AA1991.DAT:
12:	/SIDS1/qcdata/qcseq/qcseq-emb1/AA1992.DAT:
13:	/SIDS1/qcdata/qcseq/qcseq-emb1/AA1993.DAT:
14:	/SIDS1/qcdata/qcseq/qcseq-emb1/AA1994.DAT:
15:	/SIDS1/qcdata/qcseq/qcseq-emb1/AA1995.DAT:
16:	/SIDS1/qcdata/qcseq/qcseq-emb1/AA1996.DAT:
17:	/SIDS1/qcdata/qcseq/qcseq-emb1/AA1997.DAT:
18:	/SIDS1/qcdata/qcseq/qcseq-emb1/AA1998.DAT:
19:	/SIDS1/qcdata/qcseq/qcseq-emb1/AA1999.DAT:
20:	/SIDS1/qcdata/qcseq/qcseq-emb1/AA2000.DAT:
21:	/SIDS1/qcdata/qcseq/qcseq-emb1/AA2001.DAT:
22:	/SIDS1/qcdata/qcseq/qcseq-emb1/AA2002.DAT:
23:	/SIDS1/qcdata/qcseq/qcseq-emb1/AA2003.DAT:
24:	/SIDS1/qcdata/qcseq/qcseq-emb1/AA2004.DAT:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DP	ID	Description
1	797	100.0	150	20	AAV42851	Chimeric protein
2	555.5	69.7	107	20	AAV42860	CSH-miniprotein
3	470	59.0	92	20	AAV42856	Human growth hormo
4	470	59.0	134	20	AAV92265	Human anti-angioge
5	470	59.0	191	23	ABG94861	Human growth hormo
6	466	58.5	172	20	AAV91299	Human growth hormo
7	466	58.5	132	20	AAV92264	Human anti-angioge
8	465	58.3	140	10	AAV91041	Human growth hormo
9	465	58.3	261	10	AAV91299	Human growth hormo

10	465	58.3	252	12	AAV11740	Human growth hormo
11	465	58.3	310	11	AAV03255	Fusion protein of
12	464	58.2	191	23	ABG31862	Mature human growt
13	463	58.1	191	23	ABG94860	Human growth hormo
14	463	58.1	191	23	ABG94977	Human growth hormo
15	462	58.0	144	11	AAV03313	Segment of B-cell
16	462	58.0	191	23	ABG94975	Human growth hormo
17	462	58.0	191	23	ABG94976	Human growth hormo
18	462	58.0	262	7	AAV61033	Human beta-nerve g
19	461	57.8	191	18	AAV20110	Protein sequence o
20	461	57.8	191	20	AAV04396	Natural human 22kD
21	461	57.8	191	21	AAV78425	Human growth hormo
22	461	57.8	191	22	AAV17485	Human growth hormo
23	461	57.8	191	22	AAV17486	Human growth hormo
24	461	57.8	191	23	ABG94863	Human growth hormo
25	461	57.8	191	23	ABG94864	Human growth hormo
26	461	57.8	191	23	ABG94907	Human growth hormo
27	461	57.8	191	23	ABG94908	Human growth hormo
28	461	57.8	191	23	ABG94909	Human growth hormo
29	461	57.8	191	23	ABG94910	Human growth hormo
30	461	57.8	191	23	ABG94911	Human growth hormo
31	461	57.8	191	23	ABG94912	Human growth hormo
32	461	57.8	191	23	ABG94913	Human growth hormo
33	461	57.8	191	23	ABG94914	Human growth hormo
34	461	57.8	191	23	ABG94915	Human growth hormo
35	461	57.8	191	23	ABG94916	Human growth hormo
36	461	57.8	191	23	ABG94917	Human growth hormo
37	461	57.8	191	23	ABG94918	Human growth hormo
38	461	57.8	191	23	ABG94919	Human growth hormo
39	461	57.8	191	23	ABG94920	Human growth hormo
40	461	57.8	191	23	ABG94921	Human growth hormo
41	461	57.8	191	23	ABG94922	Human growth hormo
42	461	57.8	191	23	ABG94923	Human growth hormo
43	461	57.8	191	23	ABG94924	Human growth hormo
44	461	57.8	191	23	ABG94925	Human growth hormo
45	461	57.8	191	23	ABG94926	Human growth hormo

ALIGNMENTS

RESULT 1
ID: AAV42861 standard: protein: 150 AA.
XX AAV42861:
AC AAV42861:
DI 19-JAN-2000 (first entry)
XX Chimeric protein, SEQ ID 7.
DF Chimeric protein, SEQ ID 7.
XX
XX Insulin; precursor: growth hormone; chaperone; intramolecular;
KW folding; conformation; chimeric protein; cleavable; recombinant;
KW production; yield.
XX
XX Synthetic.
CS Homo sapiens.
XX
XX WGS050302-AL.
XX
XX 07-OCT-1999.
XX
XX 31-KAR-1998; 98WO-CN00052.
XX
XX 31-KAR-1998; 98WO-CN00052.
PA (TONG-) TONGHUA GANTECH BIOTECHNOLOGY LTD.
XX
XX Can 2;
XX
XX WFI: 1999-610839/52.
XX
XX New chimeric proteins containing human growth hormone fragment, used

PT particularly for the production of human insulin

Claim 14; Page 40-51; 46pp; English.

XX This sequence represents a chimeric protein, which contains an
XX N-terminal fragment of human growth hormone (hGH) of the sequence given
XX in AAY42856, a cleavable peptide linker (AAY42857), and a human insulin
XX precursor comprising insulin A and B chains (AAY42859). The hGH portion
XX of the chimeric protein acts as an intramolecular chaperone (IMC) for
XX the insulin precursor, enabling it to fold correctly. The cleavable
XX peptide linker has a C-terminal Arg residue which enables the hGH portion
XX of the chimeric protein to be removed after folding has taken place.
XX Production of recombinant human insulin via an hGH-proinsulin chimeric
XX protein can provide human insulin with correctly linked cysteine bridges
XX with fewer necessary procedural steps, and hence resulting in a higher
XX yield of human insulin. The IMC sequences not only protect insulin
XX sequences from intracellular degradation by a microorganism host, but
XX also promote the folding of the fused insulin precursor, facilitate the
XX solubility of the fusion protein and decrease the intermolecular
XX interactions among the fusion proteins, thus allowing folding of the
XX fused insulin precursor at commercially useful high concentrations. The
XX procedural steps of cyanogen bromide cleavage, oxidative sulphytolysis
XX and related purification steps can thus be eliminated, along with the use
XX of high concentrations of mercaptan or the use of hydrophobic absorbent
XX resins.

XX Sequence 150 AA;

Query Match 100.0% Score 797; DB 20; Length 150;
Best Local Similarity 100.0% Pred. No. 10-42;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFPTPLSRLEFDNMLRAHRLHQLAFDTYQEPFAYIPKEQKYSFLQNPQTSLSPSESIP 60
DB 1 MFPTPLSRLEFDNMLRAHRLHQLAFDTYQEPFAYIPKEQKYSFLQNPQTSLSPSESIP 60
QY 61 TPSNREETOQKSNLELRISLLLIQSWLGPVLGTGPRFVNHQHLGCSHLVEALYLVCGER 120
DB 61 TPSNREETOQKSNLELRISLLLIQSWLGPVLGTGPRFVNHQHLGCSHLVEALYLVCGER 120
QY 121 GFFYTPKTRGIVEQCCISCSLYQLENYCN 150
DB 121 GFFYTPKTRGIVEQCCISCSLYQLENYCN 150

RESULT 2

AAY42860

ID AAY42860 standard; protein; 107 AA.

XX AAY42860;

CT 19-JAN-2000 (first entry)

XX hGH-mini proinsulin chimeric protein.

XX insulin; precursor; growth hormone; chaperone; intramolecular;
XX folding; conformation; chimeric protein; cleavable; recombinant;
XX production; yield.

XX Synthetic.

XX Homo sapiens.

XX WO950302-A1.

XX 07-OCT-1999.

XX 31-MAR-1998; 98WO-CN00052

XX 31-MAR-1998; 98WO-CN00052

XX (TONG-) TONGHUA GANTECH BIO-TECHNOLOGY LTD.

XX Can 2;

XX WPI; 1993-610839/52.

XX New chimeric proteins containing human growth hormone fragment, used
XX particularly for the production of human insulin

XX Claim 13; Page 30; 46pp; English.

XX This sequence represents a chimeric protein, hGH-mini proinsulin
XX This chimeric protein contains an N-terminal fragment of human growth
XX hormone (hGH) of the sequence given in AAY42855, a cleavable peptide
XX linker (AAY42857), and a human insulin precursor comprising insulin
XX A and B chains (AAY42859). The hGH portion of the chimeric protein acts
XX as an intramolecular chaperone (IMC) for the insulin precursor,
XX enabling it to fold correctly. The cleavable peptide linker has a
XX C-terminal Arg residue which enables the hGH portion of the
XX chimeric protein to be removed after folding has taken place. Production
XX of recombinant human insulin via an hGH-proinsulin chimeric protein can
XX provide human insulin with correctly linked cysteine bridges with
XX fewer necessary procedural steps, and hence resulting in a higher yield
XX of human insulin. The IMC sequences not only protect insulin sequences
XX from intracellular degradation by a microorganism host, but also promote
XX the folding of the fused insulin precursor, facilitate the solubility of
XX the fusion protein and decrease the intermolecular interactions among
XX the fusion proteins, thus allowing folding of the fused insulin precursor
XX at commercially useful high concentrations. The procedural steps of
XX cyanogen bromide cleavage, oxidative sulphytolysis and related
XX purification steps can thus be eliminated, along with the use of high
XX concentrations of mercaptan or the use of hydrophobic absorbent resins.

XX Sequence 107 AA;

Query Match 69.7% Score 555.5; DB 20; Length 107;
Best Local Similarity 71.3% Pred. No. 8e-28;
Matches 107; Conservative 0; Mismatches 0; Indels 43; Gaps 1;

QY 1 MFPTPLSRLEFDNMLRAHRLHQLAFDTYQEPFAYIPKEQKYSFLQNPQTSLSPSESIP 60
DB 1 MFPTPLSRLEFDNMLRAHRLHQLAFDTYQEPFAYIPKEQKYSFLQNPQTSLSPSESIP 60
QY 61 TPSNREETOQKSNLELRISLLLIQSWLGPVLGTGPRFVNHQHLGCSHLVEALYLVCGER 120
DB 61 TPSNREETOQKSNLELRISLLLIQSWLGPVLGTGPRFVNHQHLGCSHLVEALYLVCGER 120
QY 121 GFFYTPKTRGIVEQCCISCSLYQLENYCN 150
DB 121 GFFYTPKTRGIVEQCCISCSLYQLENYCN 150

RESULT 5

AAY42856

ID AAY42856 standard; protein; 92 AA.

XX AAY42856;

CT 19-JAN-2000 (first entry)

XX Human growth hormone (hGH) N-terminal fragment #2.

XX Growth hormone; chaperone; intramolecular; insulin; precursor;
XX folding; conformation; chimeric protein; cleavable; recombinant;
XX production; yield.

XX Homo sapiens.

XX WO950302-A1.

XX 07-OCT-1999.

XX 31-MAR-1998; 98WO-CN00052.

XX 31-MAR-1998; 98WO-CN00052.

(TONGH) TONGHUA GANFENG BIOTECHNOLOGY LTD.
 Gan Z;
 WPI: 1999-510839/52.
 New chimeric proteins containing human growth hormone fragment, used particularly for the production of human insulin.
 Claim 5: Page 28; 46pp; English.
 This sequence represents an N-terminal fragment of human growth hormone (hGH) which is a component of a chimeric protein (AA42861) which also contains a human insulin precursor (AA42855). The hGH portion of the chimeric protein acts as an intramolecular chaperone (IMC) for the insulin precursor, enabling it to fold correctly. A cleavable peptide linker with a C-terminal Arg residue (AA42857) enables the hGH portion of the chimeric protein to be removed after folding has taken place. Production of recombinant human insulin via an hGH-proinsulin chimera can provide human insulin with correctly linked cysteine bridges with fewer necessary procedural steps, and hence resulting in a higher yield of human insulin. The IMC sequences not only protect insulin sequences from intracellular degradation by a microorganism host, but also promote the folding of the fused insulin precursor, facilitate the solubility of the fusion protein and decrease the intermolecular interactions among the fusion proteins, thus allowing folding of the fused insulin precursor at commercially useful high concentrations. The procedural steps of cyanogen bromide cleavage, oxidative sulphhydryls and related purification steps can thus be eliminated, along with the use of high concentrations of mercaptan or the use of hydrophobic absorbent resins.
 Sequence 92 AA:
 Query Match 59.0%; Score 470; DB 20; Length 134;
 Best Local Similarity 100.0%; Pred. No. 150-22;
 Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MEPTIPSLRFDNAMLRAHRLHQLAFTDYQFEFAYIPKQKYSFLONPQISLFSSESIP 60
 DB 1 MEPTIPSLRFDNAMLRAHRLHQLAFTDYQFEFAYIPKQKYSFLONPQISLFSSESIP 60
 QY 61 TPSNREFTQOKSNLELLRISILLIOSWLEPVQ 92
 DB 61 TPSNREFTQOKSNLELLRISILLIOSWLEPVQ 92
 RESULT 4
 AAW92265 standard; Protein: 134 AA.
 AAW92265;
 08-JUN-1999 (first entry)
 Human anti-angiogenic peptide 16K hGH Met-1-Pro133.
 Human; anti-angiogenic; prolactin; placental lactogen; hPL; angiogenesis; growth hormone; hGH; hGH-V; capillary endothelial cell proliferation; placental vascularisation; pregnancy; treatment; angiogenic disease; tumour; inhibitor; malignant; angiofibroma; arteriovenous malformation; arthritis; atherosclerotic plaques; corneal graft neovascularisation; wound healing; proliferative retinopathy; macular degeneration; trachoma; granulation; glaucoma; ocular; uveitis; fracture; Osler-Weber syndrome; psoriasis; fibroplasia; scleroderma; Kaposi's sarcoma; vascular adhesion; ulcer; leukaemia; reproductive disorder; contraceptive agent; gene therapy; pre-eclampsia; intrauterine growth retardation; placental dysfunction.
 Homo sapiens.
 WO9851323-A1.

15-NOV-1998.
 12-MAY-1998; 98WO-US09591.
 13-MAY-1997; 97US-0046394.
 (RECC) UNIV CALIFORNIA.
 Martial JA, Struman J, Taylor R, Weiner RI.
 WPI: 1999-045192/04.
 N-PSDB; AAX01707.
 New anti-angiogenic peptides - comprise N-terminal fragments of human placental lactogen, human growth hormone, growth hormone variant or human prolactin
 Claim 4: Page 49-50; 87pp; English.
 This invention describes novel human anti-angiogenic peptides derived from 10 to 150 consecutive amino acids selected from the N-terminal end of human placental lactogen (hPL), human growth hormone (hGH), growth hormone variant (hGH-V), or human prolactin. Such peptides (i) inhibit capillary endothelial cell proliferation and organisation (ii) inhibit angiogenesis in chick chorioallantoic membrane and (iii) binds to at least one specific receptor which does not bind an intact full length hGH, hPL, prolactin or hGH-V. The invention also describes a method for diagnosing a probable abnormality of placental vascularisation during pregnancy. The peptides can be used for treating an angiogenic disease in a subject, for inhibiting tumour formation or growth in a patient or for modulating vascularisation of a patient's placenta. In particular, the peptides can be used for preventing or treating e.g. malignant tumours, angiofibroma, arteriovenous malformation, arthritic such as rheumatoid arthritis, atherosclerotic plaques, corneal graft neovascularisation, delayed wound healing, proliferative retinopathy such as diabetic retinopathy, macular degeneration, granulations such as those occurring in haemophilic joints, inappropriate vascularisation in wound healing such as hypertrophic scars or keloid scars, neovascular glaucoma, ocular tumour, uveitis, non-union fractures, Osler-Weber syndrome, psoriasis, pyogenic glaucoma, retrolental fibroplasia, scleroderma, solid tumours, Kaposi's sarcoma, trachoma, vascular adhesions, chronic varicose ulcers, leukaemia, and reproductive disorders such as follicular and luteal cysts and choriocarcinoma. They can also be used as contraceptive agents. DNA encoding the peptides can be used in gene therapy. The measurement of abnormal levels of N-terminal fragments of hGH, hGH-V, prolactin or hPL can be used in assays for impairment of vascular development associated with pre-eclampsia, intrauterine growth retardation, and placental dysfunction.
 Sequence 134 AA:
 Query Match 59.0%; Score 470; DB 20; Length 134;
 Best Local Similarity 100.0%; Pred. No. 20-22;
 Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MEPTIPSLRFDNAMLRAHRLHQLAFTDYQFEFAYIPKQKYSFLONPQISLFSSESIP 60
 DB 1 MEPTIPSLRFDNAMLRAHRLHQLAFTDYQFEFAYIPKQKYSFLONPQISLFSSESIP 60
 QY 61 TPSNREFTQOKSNLELLRISILLIOSWLEPVQ 92
 DB 61 TPSNREFTQOKSNLELLRISILLIOSWLEPVQ 92
 RESULT 5
 AAG54861 standard; Protein: 191 AA.
 AAG54861;
 03-DEC-2002 (first entry)
 Human growth hormone mutant hHPL (111-129).

XX Growth hormone; placental lactogen; prolactin; active domain; hGH;
KW structure-function relationship; segment-substituted polypeptide;
KW mutant; mutin.

XX Homo sapiens.
OS Synthetic.

XX US6428954-B1.

XX 06-AUG-2002.

XX 06-JUN-1995; 95US-0481039

XX 26-OCT-1992; 53US-0438066

XX 27-APR-1992; 92US-0575204

XX 13-OCT-1992; 92US-0960227

XX 02-FEB-1994; 94US-0190723

XX 28-OCT-1988; 88US-0264511

XX (GETH) GENENTECH INC.

XX Wells JA, Cunningham BC;

XX WPI; 2002-696875/75.

XX Identifying active domains within cloned polypeptides of known amino
PT acid sequence by substituting analog segments into the parent
PT polypeptide is useful to determine the relationship between structure
PT and function.

PS Example 1; Page -; 86pp; English.

XX The invention relates to identifying an unknown active domain in a region
CC of known amino acid sequence in a parent polypeptide e.g. human growth
CC hormone (hGH) which has been cloned and has a pre-identified biological
CC activity, where the active domain interacts with a target when the parent
CC polypeptide is in its native-folded form and the interaction is
CC responsible for the biological activity comprising: (a) comparing the
CC amino acid sequence or polypeptide structure in the region of known amino
CC acid sequence of hGH with the amino acid sequence or polypeptide
CC structure in a region of known amino acid sequence of an analogue
CC polypeptide (e.g. prolactin, placental lactogen or porcine growth
CC hormone) which has at least 1% homology with hGH alpha-carbon
CC coordinates within about 2.5-5 angstroms of hGH alpha-carbon
CC for about 50% of the analogue sequence, where any interaction of the
CC analogue with the target is different from target interaction with hGH;
CC (b) substituting DNA encoding an analogous polypeptide segment from the
CC analogue into DNA encoding the full length hGH, and expressing a
CC segment-substituted polypeptide; (c) contacting the segment-substituted
CC polypeptide with the target to determine interaction; (d) repeating steps
CC (b) and (c) with a second analogous polypeptide segment; and
CC (e) comparing the difference between activity of the first and second
CC segment-substituted polypeptides as an indication of the location of
CC the unknown active domain in hGH. The method is useful for determination
CC the relationship between structure and function of known polypeptide
CC sequences. The present sequence is that of human growth hormone
CC mutant substituted with residues from an hGH analogue (prolactin,
CC placental lactogen or porcine growth hormone).
CC Note: The present sequence is not shown in the specification but was
CC created by the indexer using the mature hGH sequence and information
CC contained in the specification.

XX Sequence 191 AA;

Query Match 59.0%; Score 479; DB 23; Length 191;

Best Local Similarity 69.2%; Pred. No. 2, 6e 22;

Matches 10; Conservative 8; Mismatches 19; Indels 19; Gaps 4;

QY 2 FTTPLSLRFDNAMLRAHRLHQLAFDTYQFEFAYIPKEQKYSFLQNTQTSILCFSESIP 61

DB 1 FTTPLSLRFDNAMLRAHRLHQLAFDTYQFEFAYIPKEQKYSFLQNTQTSILCFSESIP 61

QY 62 PSNRSEETQOKSNLELLRLISLLIQSNLWFVQIGTGRFVNQHLGSG-----H 108
DB 62 PSNRSEETQOKSNLELLRLISLLIQSNLWFVQIGTGRFVNQHLGSG-----H 108
QY 109 LVEALYLVGGERGFFTPKTRGIVEQ 134
DB 120 LVEALYLVGGERGFFTPKTRGIVEQ 134
QY 120 LVEALYLVGGERGFFTPKTRGIVEQ 134
DB 120 LVEALYLVGGERGFFTPKTRGIVEQ 134

RESULT 6

AAAP90125

XX AAAP90125 standard; protein; 192 AA.

XX AAAP90125

XX 25-MAR-2003 (updated)

XX 06-FEB-1996 (revised)

XX 02-NOV-1989 (first entry)

XX Human growth hormone.

XX Human growth hormone; fusion protein; recombinant

XX Human growth hormone; fusion protein; recombinant

XX Homo sapiens (Human).

XX JF01144941-A.

XX 07-JUN-1989.

XX 02-DEC-1987; 87JP-0304937.

XX 02-DEC-1987; 87JP-0304937.

XX (WAKT) WAKUNAGA SEIYAKU KK.

XX WPI; 1989-209284/29.

XX N-PSDB; AAN90269.

XX Recombinant vector contg. fusion protein - consisting of human

XX growth hormone or deriv. ligated to foreign protein, for stability

XX and high yield.

XX Disclosure; Fig 1; 19pp; Japanese.

XX The invention consists of a vector contg. a fusion protein which is

XX formed by ligating, downstream of a promoter, hGH or a deriv. (pref.

XX formed by substn. of Met-14 with Leu) and a foreign protein.

XX Stability of the vector in the host is greatly increased so the

XX protein yield is higher.

XX (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 192 AA;

Query Match 58.5%; Score 466; DB 10; Length 192;

Best Local Similarity 70.5%; Pred. No. 4, 6e 22;

Matches 10; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 1 MFTTPLSLRFDNAMLRAHRLHQLAFDTYQFEFAYIPKEQKYSFLQNTQTSILCFSESIP 60

DB 1 MFTTPLSLRFDNAMLRAHRLHQLAFDTYQFEFAYIPKEQKYSFLQNTQTSILCFSESIP 60

QY 61 TPSNRSEETQOKSNLELLRLISLLIQSNLWFVQIGTGRFVNQHLGSG-----HLV 110

DB 61 TPSNRSEETQOKSNLELLRLISLLIQSNLWFVQIGTGRFVNQHLGSG-----HLV 110

QY 111 EALYLVGGERGFFTPKTRGIVEQ 134

DB 120 EGIOTLNGHLEDG --SPRTGQIFKQ 142

RESULT 7

AAN92254

AAW92264 Standard; Protein; 192 AA.
 AAW92264:
 08-JUN-1999 (first entry)
 Human anti-angiogenic peptide hGH Met-1Phc191.
 Human; anti-angiogenic; prolactin; placental lactogen; hPL; angiogenesis;
 growth hormone; hGH; hGH-V; capillary endothelial cell proliferation;
 placental vasculature; pregnancy; treatment; angiogenic disease;
 tumor; inhibitor; malignancy; angiofibroma; arteriovenous malformation;
 arthritis; atherosclerotic plaques; corneal graft neovascularization;
 wound healing; proliferative retinopathy; macular degeneration; trachoma;
 granulation; glaucoma; ocular; uveitis; fracture; Oster-Weber syndrome;
 psoriasis; fibroplasia; scleroderma; Kaposi's sarcoma; vascular adhesion;
 ulcer; leukemia; reproductive disorder; contraceptive agent;
 gene therapy; pre-eclampsia; intrauterine growth retardation;
 placental dysfunction.
 OS Homo sapiens.
 XX
 XX W09851323-A1.
 XX
 XX 13-NOV-1998.
 XX
 XX 12-MAY-1998; 98W0-US03691.
 XX
 XX 13-MAY-1997; 97US-0043394.
 XX
 XX (PACC) CNIV CALIFORNIA.
 XX
 XX Martial JA, Struman I, Taylor R, Weiner M;
 WPI: 1999-045122/0A.
 DR N-PSDB; AAW92706.
 XX
 XX New anti-angiogenic peptides - comprise N-terminal fragments of
 PT human placental lactogen; human growth hormone; growth hormone
 PT variant or human prolactin.
 XX
 PS Example 3: Page 49; 87pp; English.
 XX
 CC This invention describes novel human anti-angiogenic peptides derived
 CC from 10 to 150 consecutive amino acids selected from the N-terminal end
 CC of human placental lactogen (hPL), human growth hormone (hGH), growth
 CC hormone variant (hGH-V), or human prolactin. Such peptides (i) inhibit
 CC capillary endothelial cell proliferation and organisation (ii) inhibit
 CC angiogenesis in chick chorioallantoic membrane and (iii) binds to at
 CC least one specific receptor which does not bind an intact full length
 CC hGH, hPL, prolactin or hGH-V. The invention also describes a method for
 CC diagnosing a probable abnormality of placental vascularisation during
 CC pregnancy. The peptides can be used for treating an angiogenic disease in
 CC a subject, for inhibiting tumor formation or growth in a patient or for
 CC modulating vascularisation of a patient's placenta. In particular, the
 CC peptides can be used for preventing or treating e.g. malignant tumors,
 CC angiofibroma, arteriovenous malformation, arthritic such as rheumatoid
 CC arthritis, atherosclerotic plaques, corneal graft neovascularisation,
 CC delayed wound healing, proliferative retinopathy such as diabetic
 CC retinopathy, macular degeneration, granulation such as those occurring
 CC in haemophilic joints, inappropriate vascularisation in wound healing
 CC such as hypertrophic scars or keloid scars, neovascular glaucoma, ocular
 CC tumor, uveitis, non-union fractures, Oster-Weber syndrome, psoriasis,
 CC pyogenic glaucoma, retrolental fibroplasia, scleroderma, solid tumors,
 CC Kaposi's sarcoma, trachoma, vascular adhesions, chronic varicose ulcers,
 CC leukemia, and reproductive disorders such as follicular and luteal cysts
 CC and choriocarcinoma. They can also be used as contraceptive agents. DNA
 CC encoding the peptides can be used in gene therapy. The measurement of
 CC abnormal levels of N-terminal fragments of hGH, hGH-V, prolactin or hPL
 CC can be used in assays for impairment of vascular development associated
 CC with pre-eclampsia, intrauterine growth retardation, and placental
 CC dysfunction.
 XX

SV Sequence 192 AA:
 Query Match 58.5%; Score 466; DH 20; Length 192;
 Best Local Similarity 70.5%; Pred. No. 4.6e-22;
 Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;
 QY 1 MFTTPIPLSLFDNAMLRAHRLHQLAFDTYQFEFAVLPKEQKYSFLQNPTSLSPSESIP 60
 DB 1 MEPTIPLSLRFLDNAMLAHRLHQLAFDTYQFEFAVLPKEQKYSFLQNPTSLSPSESIP 60
 QY 61 TPSNRETSQKSNLELLRISLLLIQSMLPVLQIGIGPRFVNOHLCGS-----HLV 110
 DB 61 TPSNRETSQKSNLELLRISLLLIQSMLPVLQIGIGPRFVNOHLCGS-----HLV 110
 QY 111 EALYLVCG--ERGFYTPKTRGVVEQ 134
 DB 120 EGIOTLMRLFDG---SPRIGQIFKQ 142
 RESULT 8
 AAW91041
 ID AAW91041 standard; protein; 140 AA.
 XX
 XX AAW91041:
 CT 14-DEC-1989 (first entry).
 DE Human growth hormone segment.
 KW Human growth hormone; fusion protein; thrombin;
 KW geriatric dementia; nervous disorders; human nerve factor.
 XX
 XX Homo sapiens (human).
 XX
 XX EP329175-A.
 XX
 XX 23-APR-1989.
 XX
 XX 17-FEB-1989; 89EP-G102795.
 XX
 XX 19-FEB-1982; 88JP-0015042.
 XX
 XX (TGT) TOSOH CORP.
 XX
 XX Chisuka E.
 XX
 XX WPI: 1989-243092/34.
 XX
 PT New human nerve growth factor gene encoding fusion protein
 PT - having cleavage site for thrombin, useful for treating geriatric
 XX dementia, etc.
 PS Disclosure: page 21; 38pp; English.
 XX
 CC Human growth hormone segment, used at the N-terminal of a fusion
 CC protein, which contains a thrombin recognition site, and human beta nerve
 CC growth factor (beta-NGF) at the C-terminal. Beta-NGF can be used to
 CC control geriatric dementia and other nervous disorders, and can be
 CC released from the fusion protein by incubation with thrombin (see
 CC AAW90577-9, AAW91034, AAW91299).
 XX
 XX Sequence 140 AA:
 Query Match 58.3%; Score 465; DH 10; Length 140;
 Best Local Similarity 98.9%; Pred. No. 4.2e-22;
 Matches 91; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MEPTIPLSLRFLDNAMLAHRLHQLAFDTYQFEFAVLPKEQKYSFLQNPTSLSPSESIP 60
 DB 1 MEPTIPLSLRFLDNAMLAHRLHQLAFDTYQFEFAVLPKEQKYSFLQNPTSLSPSESIP 60
 QY 61 TPSNRETSQKSNLELLRISLLLIQSMLPVLQIGIGPRFVNOHLCGS-----HLV 110
 DB 61 TPSNRETSQKSNLELLRISLLLIQSMLPVLQIGIGPRFVNOHLCGS-----HLV 110

PA (TOSC) TOSOM CORP.
 XX
 XX WPI: 1990-062207/09.
 DR N-PSDB: AAO02028.
 XX
 XX Prepn. of human B-cell differentiation factor - from specified DNA
 PT sequence segment. By recombinant DNA technique, gives protein of
 PI specified amino acid sequence.
 XX
 PS Claim 31: Page 9: 17pp; Japanese.
 XX
 XX The protein is produced by fusing DNA encoding BDF (1-7) with DNA
 CC encoding BDF-2 (11-5) and ligating the product into an expression vector
 CC See also AAR05311 and AAR05313.
 XX
 XX Sequence 310 AA;
 SQ

Query Match 58.3%; Score 465; DB 13; Length 400;
 Best Local Similarity 98.9%; Pred. No. 7.7e-22;
 Matches 91; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MFPTPLSRFDNKLRAFLRLHQAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSISPT 60
 Db 1 MFPTPLSRFDNKLRAFLRLHQAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSISPT 60
 QY 61 TPSNREETQCKSNLELLRLSLLQTSWLEPVQ 32
 Db 61 TPSNREETQCKSNLELLRLSLLQTSWLEPVQ 32

RESULT 12
 ABG31862
 ID ABG31862 standard; Protein: 191 AA.
 AC ABG31862;
 XX
 XX 05-NOV-2002 (first entry)
 XX
 XX Mature human growth hormone (hGH), mutant #4.
 XX
 XX Human; growth hormone; hGH; Turner's syndrome; achondroplasia;
 KW growth hormone deficiency in adults; GHDA; chronic renal insufficiency;
 KW renal failure in children; acquired immune deficiency syndrome; AIDS;
 KW AIDS wasting; cachexia; mutant; murine.
 XX
 XX Homo sapiens.
 QS Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 134
 PT /note: "W10 type Arg substituted by Lys"
 XX
 XX W0200255532-A2.
 XX
 XX 18-JUL-2002.
 XX
 XX 10-JAN-2002; 2002WO-DK00017.
 XX
 XX 11-JAN-2001; 2001DK-0950042.
 PR 11-JAN-2001; 2001US-261411P.
 XX
 XX (MAXY-) MAXYGEN APS.
 PA (MAXY-) MAXYGEN HOLDINGS LTD.
 XX
 XX Andersen KV, Drustup J, Christensen J;
 DI WPI: 2002-608345/65.
 XX
 XX New conjugates exhibiting growth hormone activity, useful for treating
 PT a disease or for manufacturing a medicament for treating a disease.
 PT e.g. Turner's syndrome, growth hormone deficiency, achondroplasia, AIDS
 PT wasting or cachexia -
 XX

IS Claim 10; Page 7: 74pp; English.
 XX
 XX The invention relates to new conjugates, which exhibit growth hormone
 CC (GH) activity and comprise at least one non-polypeptide group covalently
 CC attached to a GH polypeptide. The amino acid sequence of the conjugates
 CC differs from that of wild type human GH in at least one introduced and at
 CC least one removed amino acid residue comprising an attachment group for
 CC the first non-polypeptide group. The conjugate or pharmaceutical
 CC composition is useful for treating a disease or for manufacturing a
 CC medicament for treating a disease, e.g. Turner's syndrome, GH deficiency
 CC in adults (i.e. GHDA), achondroplasia, chronic renal insufficiency or
 CC failure (including renal failure in children), acquired immune deficiency
 CC syndrome (AIDS) wasting, cachexia in AIDS patients, or cachexia
 CC associated with other diseases. The conjugates are useful for treating a
 CC variety of disorders caused by growth hormone inadequacy. The present
 CC sequence represents the amino acid sequence of a mutant human growth
 CC hormone.
 CC Note: The present sequence is not shown in the specification but is
 CC derived from the wild type human growth hormone sequence given in SEQ ID
 CC No.2 (see ABG31857).
 XX
 XX Sequence 191 AA;
 SQ

Query Match 58.2%; Score 464; DB 23; Length 191;
 Best Local Similarity 71.0%; Pred. No. 6.1e-22;
 Matches 103; Conservative 6; Mismatches 20; Indels 16; Gaps 4;

QY 2 FPIPLSRFLFDNKLRAHRLHQAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSISPT 61
 Db 2 FPIPLSRFLFDNKLRAHRLHQAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSISPT 60
 QY 62 PSNREETQCKSNLELLRLSLLQTSWLEPVQGTGPRFVNHLCGS-----HLVE 111
 Db 62 PSNREETQCKSNLELLRLSLLQTSWLEPVQF-IRSVFANSJYVGASDSNVYLLKDLLE 119
 QY 112 ADYVQGG--ERGFEYTPKTRGIVEQ 134
 Db 120 GIOTLMGRLEDG---SPKGTGJEKQ 141

RFSLT 13
 ABG34860
 ID ABG34860 standard; Protein: 191 AA.
 AC ABG34860;
 XX
 XX 03-DEC-2002 (first entry)
 XX
 XX Human growth hormone mutant hPL (109-112).
 XX
 XX Growth hormone; placental lactogen; prolactin; active domain; hGH;
 KW structure-function relationship; segment-substituted polypeptide;
 KW mutant; murine.
 XX
 XX Homo sapiens.
 QS Synthetic.
 XX
 XX JS6428954-B-1.
 XX
 XX 06-AUG-2002.
 XX
 XX 06-JUN-1995; 95US-0483039.
 XX
 XX 26-OCT-1989; 89JS-0428066.
 PR 27-APR-1992; 92JS-0875204.
 PR 13-OCT-1992; 92JS-0960227.
 PR 02-FEB-1994; 94US-0190723.
 PR 28-OCT-1988; 88US-0264611.
 XX
 XX (GETH) GENENTECH INC.
 PA Wells JA, Cunningham BC;
 XX
 XX

DR WPI: 2002-696875/75.

XX Identifying active domains within cloned polypeptides of known amino

PT acid sequence by substituting analog segments into the parent

PT polypeptide is useful to determine the relationship between structure

PT and function.

XX Example 1: Page -: 86pp; English.

PS The invention relates to identifying an unknown active domain in a region

XX of known amino acid sequence in a parent polypeptide e.g. human growth

CC hormone (hGH) which has been cloned and has a pre-identified biological

CC activity, where the active domain interacts with a target when the parent

CC polypeptide is in its native-folded form and the interaction is

CC responsible for the biological activity comprising: (a) comparing the

CC amino acid sequence of polypeptide structure in the region of known amino

CC acid sequence of hGH with the amino acid sequence of an analogue

CC analogue with the target is different from target interaction with hGH;

CC (b) substituting DNA encoding an analogous polypeptide segment from the

CC analogue into DNA encoding the full length hGH, and expressing a

CC segment-substituted polypeptide; (c) contacting the segment-substituted

CC polypeptide with the target to determine interaction; (d) repeating steps

CC (b) and (c) with a second analogous polypeptide segment; and

CC (e) comparing the difference between activity of the first and second

CC segment-substituted polypeptides as an indication of the location of

CC the unknown active domain in hGH. The method is useful for determining

CC the relationship between structure and function of known polypeptide

CC sequences. The present sequence is that of human growth hormone

CC mutant substituted at functionally important residues and used in

CC the method of the invention.

CC Note: The present sequence is not shown in the specification but was

CC created by the indexer using the mature hGH sequence and information

CC contained in the specification.

XX Sequence 191 AA:

50 Query Match 58.1%; Score 453; DB 23; Length 191;

Best Local Similarity 71.0%; Pred. No. 7.e-22;

Matches 103; Conservative 3; Mismatches 16; Indels 16; Gaps 5;

QY 2 FPTPLSLRLFNAMLRHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNPQTSLSFSESPT 61

DB 1 FPTPLSLRLFNAMLRHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNPQTSLSFSESPT 60

QY 62 PSNREETQCKSNLELLRLSLILIQSWLEPVOIGTPREVNHLCGS-----HLV-----E 111

DB 61 PSNREETQCKSNLELLRLSLILIQSWLEPVOF-LRSVPANSLVAGASDNYVLLKOLEE 119

QY 112 ALYLYCG--ERGFYTPKIRGIVEQ 134

DB 120 GLOTLMGRLEDS---SVRTSGIRKQ 141

RESULT 14

ABG94977

ID ABG94977 standard: Protein; 191 AA.

XX AC ABG94977;

XX CT 03-DEC-2002 (first entry)

XX DE Human growth hormone mutant with 3A.

XX KW Growth hormone; placental lactogen; prolactin; active domain; hGH;

XX KW structure-function relationship; segment-substituted polypeptides;

XX KW mutant; mutain.

OS Homo sapiens.

Synthetic.

XX US6428954-B1.

XX 06-AUG-2002.

XX 06-JUN-1995; 9505 0483033.

XX 26-OCT-1989; 8905-0426066.

XX 27-APR-1992; 923S-0852234.

XX 13-OCT-1992; 923S-0960227.

XX 02-FEB-1994; 9405-0190723.

XX 28-OCT-1988; 8805-0264612.

XX (GETH) GENENTECH INC.

XX Wells JA. Cunningham BC;

XX WPI: 2002-696875/75.

PT Identifying active domains within cloned polypeptides of known amino

PT acid sequence by substituting analog segments into the parent

PT polypeptide is useful to determine the relationship between structure

PT and function.

XX Example 16: Page -: 86pp; English.

PS The invention relates to identifying an unknown active domain in a region

XX of known amino acid sequence in a parent polypeptide e.g. human growth

CC hormone (hGH) which has been cloned and has a pre-identified biological

CC activity, where the active domain interacts with a target when the parent

CC polypeptide is in its native-folded form and the interaction is

CC responsible for the biological activity comprising: (a) comparing the

CC amino acid sequence of polypeptide structure in the region of known amino

CC acid sequence of hGH with the amino acid sequence of an analogue

CC structure in a region of known amino acid sequence of an analogue

CC polypeptide (e.g. prolactin, placental lactogen or porcine growth

CC hormone) which has at least 15% homology with hGH alpha-carbon

CC coordinates within about 2-3.5 angstroms of hGH alpha-carbon

CC for about 60% of the analogue sequence, where any interaction of the

CC analogue with the target is different from target interaction with hGH;

CC (b) substituting DNA encoding an analogous polypeptide segment from the

CC analogue into DNA encoding the full length hGH, and expressing a

CC segment-substituted polypeptide; (c) contacting the segment-substituted

CC polypeptide with the target to determine interaction; (d) repeating steps

CC (b) and (c) with a second analogous polypeptide segment; and

CC (e) comparing the difference between activity of the first and second

CC segment-substituted polypeptides as an indication of the location of

CC the unknown active domain in hGH. The method is useful for determining

CC the relationship between structure and function of known polypeptide

CC sequences. The present sequence is that of human growth hormone

CC mutant substituted at functionally important residues and used in

CC the method of the invention.

CC Note: The present sequence is not shown in the specification but was

CC created by the indexer using the mature hGH sequence and information

CC contained in the specification.

XX Sequence 191 AA:

50 Query Match 58.1%; Score 463; DB 23; Length 191;

Best Local Similarity 70.3%; Pred. No. 7.le 22;

Matches 104; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 2 FPTPLSLRLFNAMLRHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNPQTSLSFSESPT 61

DB 1 FPTPLSLRLFNAMLRHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNPQTSLSFSESPT 60

QY 62 PSNREETQCKSNLELLRLSLILIQSWLEPVOIGTPREVNHLCGS-----HLVE 111

DB 61 PSNREETQCKSNLELLRLSLILIQSWLEPVOF-LASVPANSLVAGASDNYVLLKOLEE 119

QY 112 ALYLYCG--ERGFYTPKIRGIVEQ 134

```

DB      120 GIQILMGLEDG--SPNIGQIFKQ 141

RESULT 15
AAR05313
ID      AAR05313 standard: protein; 144 AA.
XX
AC      AAR05313;
XX
DT      19-JUL-1990 (first entry)
XX
DE      Segment of B-cell stimulatory factor-2 (12-5).
XX
KW      B-cell stimulatory factor-2; interleukin-5.
XX
OS      Homo sapiens.
XX
PN      JP02013375-A.
XX
PD      17-JUN-1990.
XX
PF      01-JUL-1988; 8BJP-0162556
XX
PR      01-JUL-1988; 8BJP-0162556
XX
PA      (TOX) ) TOSOH CORP.
XX
DR      WPI; 1990-062207/39.
DR      N-PSDB; AACC0208.
XX
XX      Prep. of human B-cell differentiation factor - from specified DNA
PT      sequence segment, by recombinant DNA technique, gives protein of
PT      specified amino acid sequence.
XX
PS      Disclosure; Page 9; 17pp; Japanese.
XX
XX      The sequence encoding this protein can be fused with DNA encoding B-cell
CC      differentiation factor (IL-6) and ligated into an expression vector for
CC      prodn. of a fusion protein.
CC      See also AAR05311.
XX
SQ      Sequence 14; AA.

Query Match 56.6%; Score 452; DB 11; Length 144;
Best Local Similarity 87.8%; Pred. No. 6,60-221;
Matches 93; Conservative 1; Mismatched 1; Indels 0; Gaps 0;

QY      1 MEPTIFLSKLPNAMLRAHRLHQLAFETYQFEFEAYIFKQKYSFLQNTSLSESSEIF 60
        IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
DB      1 MEPTIFLSKLPNAMLRAHRLHQLAFETYQFEFEAYIFKQKYSFLQNTSLSESSEIF 60
        IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
QY      61 TPSNREETQOKSNLELRISLLIGSNLEPVQ 72
        IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
DB      61 TPSNREETQOKSNLELRISLLIGSNLEPVQ 92
        IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII

```

Search completed: September 15, 2003, 12:38:27
 Job time : 45.358 secs

Result No.	Score	Query Match	Length	ID	Description
1	466	58.5	192	US-08-093-384-1	Sequence 1, Appl
2	462	57.8	191	US-08-284-528-5	Sequence 1, Appl
3	462	57.8	194	US-08-381-521-4	Sequence 4, Appl
4	461	57.8	194	US-08-450-526-4	Sequence 4, Appl
5	461	57.8	217	US-08-589-528-10	Sequence 10, Appl
6	461	57.8	217	US-08-784-582-19	Sequence 19, Appl
7	461	57.8	217	US-08-785-273-15	Sequence 15, Appl
8	461	57.8	217	US-08-755-626-11	Sequence 11, Appl
9	461	57.8	217	US-09-284-876-1	Sequence 1, Appl
10	461	57.8	241	US-09-424-6305-25	Sequence 25, Appl
11	461	57.8	243	US-09-230-070-66	Sequence 66, Appl
12	461	57.8	274	US-08-784-582-73	Sequence 73, Appl
13	461	57.8	363	US-08-784-582-73	Sequence 73, Appl
14	455	57.1	191	US-09-465-461-1	Sequence 1, Appl
15	455	57.1	217	US-08-187-756-4	Sequence 4, Appl
16	455	57.1	217	US-08-469-486-51	Sequence 51, Appl
17	455	57.1	217	US-08-459-653-52	Sequence 52, Appl
18	455	57.1	217	US-08-710-324-4	Sequence 4, Appl
19	455	57.1	217	US-09-421-657-4	Sequence 4, Appl
20	454	57.0	400	US-08-420-819-37	Sequence 37, Appl
21	454	57.0	401	US-08-420-819-36	Sequence 36, Appl
22	448	56.2	191	US-08-800-216C-18	Sequence 18, Appl
23	446	56.0	191	US-08-800-215C-16	Sequence 16, Appl
24	446	56.0	191	US-08-800-215C-20	Sequence 20, Appl
25	365.5	45.9	176	US-08-791-728-1	Sequence 1, Appl
26	365.5	45.9	176	US-08-990-774-1	Sequence 1, Appl
27	359.5	45.1	176	US-08-791-728-2	Sequence 2, Appl

```

1  TOPOLOGY: 112cat
2  US-08-093-383-1
3
4  Query Match      58.5%  Score 461: DB 4: Length 191:
5  Best Local Similarity 70.3%  Pred. No. 3.8e-42:
6  Matches 102: Conservative 7: Mismatches 20: Indels 16: Gaps 4:
7
8  QY  2  FFTPLSLRLFDNMLRAHRLHQLAFTDYQEFEEAYIPKEQKYSFLQNPQTSLSFSSESPT 61
9  DB  3  FFTPLSLRLFDNMLRAHRLHQLAFTDYQEFEEAYIPKEQKYSFLQNPQTSLSFSSESPT 61
10
11  QY  61  TPNREFTQKSNLELRISLLIQSWLEPVOIGTGRFVNOHICGS-----HIVE 111
12  DB  61  TPNREFTQKSNLELRISLLIQSWLEPVOIGTGRFVNOHICGS-----HIVE 111
13
14  QY  112  ALYLVCG---ERGFYTPKTRGIVQ 134
15  DB  120  GIQTLMGRLDQ---SPRTGQIFKQ 144
16
17  RESULT 3
18  US-08-383-621-4
19  Sequence 4: Application US/08384621
20  Patent No. 5951972
21  GENERAL INFORMATION:
22  APPLICANT: Daley, Michael J.
23  APPLICANT: Buckwalter, Brian L.
24  APPLICANT: Cady, Susan M.
25  APPLICANT: Shieh, Hong-Ming
26  APPLICANT: Bohlen, Peter
27  APPLICANT: Seddon, Andrew P.
28  TITLE OF INVENTION: Stabilization of Somatotropins and Other
29  TITLE OF INVENTION: Proteins by Modification of Cysteine Residues
30
31  US-09-284-878-5
32  Sequence 5: Application US/09284878
33  Patent No. 6242375
34  GENERAL INFORMATION:
35  APPLICANT: Olazaran, Martha Guetero
36  APPLICANT: Saldaña, Rosa Bariera
37  APPLICANT: Salvado, Jose Maria Viader
38  TITLE OF INVENTION: Genetically Modified Methylophilic P. pastoris Yeast for the
39  TITLE OF INVENTION: Production and Secretion of the Human Growth Hormone
40  FILE REFERENCE: 1829.0016000
41  CURRENT APPLICATION NUMBER: US/09/284.878
42  PRIOR FILING DATE: 1999-07-22
43  PRIOR APPLICATION NUMBER: PCI/WX57/00033
44  PRIOR FILING DATE: 1997-10-24
45  NUMBER OF SEQ ID NOS: 9
46  SOFTWARE: PatentIn Ver. 2.1
47  SEQ ID NO 5
48  LENGTH: 191
49  TYPE: PRT
50  ORGANISM: Homo sapiens
51  US-09-284-878-5
52
53  Query Match      57.8%  Score 461: DB 4: Length 191:
54  Best Local Similarity 70.3%  Pred. No. 3.8e-42:
55  Matches 102: Conservative 7: Mismatches 20: Indels 16: Gaps 4:
56
57  QY  2  FFTPLSLRLFDNMLRAHRLHQLAFTDYQEFEEAYIPKEQKYSFLQNPQTSLSFSSESPT 61
58  DB  3  FFTPLSLRLFDNMLRAHRLHQLAFTDYQEFEEAYIPKEQKYSFLQNPQTSLSFSSESPT 61
59
60  QY  62  PSNREFTQKSNLELRISLLIQSWLEPVOIGTGRFVNOHICGS-----HIVE 111
61  DB  61  PSNREFTQKSNLELRISLLIQSWLEPVOIGTGRFVNOHICGS-----HIVE 111
62
63  QY  112  ALYLVCG---ERGFYTPKTRGIVQ 134
64  DB  120  GIQTLMGRLDQ---SPRTGQIFKQ 144
65
66  RESULT 4
67  US-08-459-906-4
68  Sequence 4: Application US/08459906
69  Patent No. 6010959
70  GENERAL INFORMATION:
71  APPLICANT: Daley, Michael J.
72  APPLICANT: Buckwalter, Brian L.
73  APPLICANT: Cady, Susan M.
74  APPLICANT: Shieh, Hong-Ming
75  APPLICANT: Bohlen, Peter
76  APPLICANT: Seddon, Andrew P.
77  TITLE OF INVENTION: Stabilization of Somatotropins and Other
78  TITLE OF INVENTION: Proteins by Modification of Cysteine Residues
79
80  US-09-423-100-7
81  Sequence 4: Application US/09423100
82  Patent No. 6242375
83  GENERAL INFORMATION:
84  APPLICANT: Daley, Michael J.
85  APPLICANT: Buckwalter, Brian L.
86  APPLICANT: Cady, Susan M.
87  APPLICANT: Shieh, Hong-Ming
88  APPLICANT: Bohlen, Peter
89  APPLICANT: Seddon, Andrew P.
90  TITLE OF INVENTION: Stabilization of Somatotropins and Other
91  TITLE OF INVENTION: Proteins by Modification of Cysteine Residues
92
93  US-09-423-100-7
94  Sequence 4: Application US/09423100
95  Patent No. 6242375
96  GENERAL INFORMATION:
97  APPLICANT: Daley, Michael J.
98  APPLICANT: Buckwalter, Brian L.
99  APPLICANT: Cady, Susan M.
100  APPLICANT: Shieh, Hong-Ming
101  APPLICANT: Bohlen, Peter
102  APPLICANT: Seddon, Andrew P.
103  TITLE OF INVENTION: Stabilization of Somatotropins and Other
104  TITLE OF INVENTION: Proteins by Modification of Cysteine Residues

```

```

1  NUMBER OF SEQUENCES: 11
2  CORRESPONDENCE ADDRESS:
3  ADDRESSEE: Dr. Estelle J. Tsevdos
4  STREET: 1937 West Main Street, P.O. Box 60
5  CITY: Stamford
6  STATE: Connecticut
7  COUNTRY: U.S.A.
8  ZIP: 06304-0060
9  COMPUTER READABLE FORM:
10  MEDIUM TYPE: Floppy disk
11  COMPUTER: IBM PC compatible
12  OPERATING SYSTEM: PC-DOS/MS-DOS
13  SOFTWARE: Patent Release #1.0, Version #1.25
14  CURRENT APPLICATION DATA:
15  APPLICATION NUMBER: US/08/383.621
16  FILING DATE: 25-FEB-1995
17  CLASSIFICATION: 514
18  PRIORITY APPLICATION DATA:
19  APPLICATION NUMBER: US 07/746,142
20  FILING DATE: 25-SEP-1991
21  ATTORNEY/AGENT INFORMATION:
22  NAME: Tsevdos, Estelle J.
23  REGISTRATION NUMBER: 31,145
24  REFERENCE/DOCKET NUMBER: 31,278-01
25  TELECOMMUNICATION INFORMATION:
26  TELEPHONE: 203-321-2756
27  TELEFAX: 203-321-2971
28  TELEX: 203-710-474-4959
29  INFORMATION FOR SEQ ID NO: 4:
30  SEQUENCE CHARACTERISTICS:
31  LENGTH: 194 amino acids
32  TYPE: amino acid
33  TOPOLOGY: linear
34  MCLEULE TYPE: protein
35  US-08-383-621-4
36
37  Query Match      57.8%  Score 461: DB 2: Length 194:
38  Best Local Similarity 70.3%  Pred. No. 3.8e-42:
39  Matches 102: Conservative 7: Mismatches 20: Indels 16: Gaps 4:
40
41  QY  2  FFTPLSLRLFDNMLRAHRLHQLAFTDYQEFEEAYIPKEQKYSFLQNPQTSLSFSSESPT 61
42  DB  4  FFTPLSLRLFDNMLRAHRLHQLAFTDYQEFEEAYIPKEQKYSFLQNPQTSLSFSSESPT 63
43
44  QY  62  PSNREFTQKSNLELRISLLIQSWLEPVOIGTGRFVNOHICGS-----HIVE 111
45  DB  64  PSNREFTQKSNLELRISLLIQSWLEPVOIGTGRFVNOHICGS-----HIVE 111
46
47  QY  112  ALYLVCG---ERGFYTPKTRGIVQ 134
48  DB  123  GIQTLMGRLDQ---SPRTGQIFKQ 144
49
50  RESULT 4
51  US-08-459-906-4
52  Sequence 4: Application US/08459906
53  Patent No. 6010959
54  GENERAL INFORMATION:
55  APPLICANT: Daley, Michael J.
56  APPLICANT: Buckwalter, Brian L.
57  APPLICANT: Cady, Susan M.
58  APPLICANT: Shieh, Hong-Ming
59  APPLICANT: Bohlen, Peter
60  APPLICANT: Seddon, Andrew P.
61  TITLE OF INVENTION: Stabilization of Somatotropins and Other
62  TITLE OF INVENTION: Proteins by Modification of Cysteine Residues
63
64  US-09-423-100-7
65  Sequence 4: Application US/09423100
66  Patent No. 6242375
67  GENERAL INFORMATION:
68  APPLICANT: Daley, Michael J.
69  APPLICANT: Buckwalter, Brian L.
70  APPLICANT: Cady, Susan M.
71  APPLICANT: Shieh, Hong-Ming
72  APPLICANT: Bohlen, Peter
73  APPLICANT: Seddon, Andrew P.
74  TITLE OF INVENTION: Stabilization of Somatotropins and Other
75  TITLE OF INVENTION: Proteins by Modification of Cysteine Residues
76
77  US-09-423-100-7
78  Sequence 4: Application US/09423100
79  Patent No. 6242375
80  GENERAL INFORMATION:
81  APPLICANT: Daley, Michael J.
82  APPLICANT: Buckwalter, Brian L.
83  APPLICANT: Cady, Susan M.
84  APPLICANT: Shieh, Hong-Ming
85  APPLICANT: Bohlen, Peter
86  APPLICANT: Seddon, Andrew P.
87  TITLE OF INVENTION: Stabilization of Somatotropins and Other
88  TITLE OF INVENTION: Proteins by Modification of Cysteine Residues

```


REFERENCE/DOCKET NUMBER: UTSD:514
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 512/418-3000
 TELEFAX: 512/474-7577
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 217 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 US-08-784-582-10

Query Match 57.8%; Score 461; DB 3; Length 217;
 Best Local Similarity 70.3%; Pred. No. 4.4e-42;
 Matches 102; Conservative 7; Mismatches 20; Gaps 4;

QY 2 FPTIPSLRFDNAKLAHRLHQIAFDYQEFREAYIPKRYKSLQNPQTSISFSESPT 61
 DB 27 FPTIPSLRFDNAKLAHRLHQIAFDYQEFREAYIPKRYKSLQNPQTSISFSESPT 86
 QY 62 PSNREETQOKSNLELRISLLIQSKLEPVQGTGPRFVNHCLGS-----HLVE 111
 DB 87 PSNREETQOKSNLELRISLLIQSKLEPVQGTGPRFVNHCLGS-----HLVE 111
 QY 112 ALYLVGQ--ERGFYTPKRGIVEQ 134
 DB 146 GIOTLMGRLEAG---SPRTGQIFKQ 167

RESULT 7
 US-08-785-271-10
 Sequence 10; Application US/08785271
 Patent No. 6194176

GENERAL INFORMATION:
 APPLICANT: Newgard, Christopher B.
 APPLICANT: Halban, Philippe A.
 APPLICANT: No. 6194176;ington, Kar: B.
 APPLICANT: Clark, Samuel A.
 APPLICANT: Thigpen, Anice E.
 APPLICANT: Kruse, Fred
 TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
 NUMBER OF SEQUENCES: 36
 CORRESPONDENCE ADDRESS:
 ADDRESS: Arnold, White & Durkee
 STREET: P.O. Box 4433
 CITY: Houston
 STATE: Texas
 COUNTRY: USA
 ZIP: 77210

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/785,271
 FILING DATE: Concurrently Herewith
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/569,028
 FILING DATE: 19-JAN-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Highlander, Steven L.
 REGISTRATION NUMBER: 37,642
 REFERENCE/DOCKET NUMBER: UTSD:513
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 512/418-3000
 TELEFAX: 512/474-7577
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 217 amino acids

TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 US-08-785-271-10

Query Match 57.8%; Score 461; DB 3; Length 217;
 Best Local Similarity 70.3%; Pred. No. 4.4e-42;
 Matches 102; Conservative 7; Mismatches 20; Gaps 4;

QY 2 FPTIPSLRFDNAKLAHRLHQIAFDYQEFREAYIPKRYKSLQNPQTSISFSESPT 61
 DB 27 FPTIPSLRFDNAKLAHRLHQIAFDYQEFREAYIPKRYKSLQNPQTSISFSESPT 86
 QY 62 PSNREETQOKSNLELRISLLIQSKLEPVQGTGPRFVNHCLGS-----HLVE 111
 DB 87 PSNREETQOKSNLELRISLLIQSKLEPVQGTGPRFVNHCLGS-----HLVE 111
 QY 112 ALYLVGQ--ERGFYTPKRGIVEQ 134
 DB 146 GIOTLMGRLEAG---SPRTGQIFKQ 167

RESULT 8

US-08-759-626-11
 Sequence 11; Application US/08759628
 Patent No. 6223446

GENERAL INFORMATION:
 APPLICANT: Altmann, Scott W.
 APPLICANT: Rock, Ferrand L.
 APPLICANT: Bazan, J. Fernando
 APPLICANT: Kastelein, Robert A.
 TITLE OF INVENTION: MUTATIONAL VARIANTS OF MAMMALIAN PROTEINS
 NUMBER OF SEQUENCES: 11
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: UNAX Research Institute
 STREET: 901 California Avenue
 CITY: Palo Alto
 STATE: California
 COUNTRY: USA
 ZIP: 94304-1104

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/759,628
 FILING DATE: 05-DEC-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/008,574
 FILING DATE: 06-DEC-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Ching, Edwin P.
 REGISTRATION NUMBER: 34,090
 REFERENCE/DOCKET NUMBER: DX05520
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-852-9196
 TELEFAX: 415-496-1200
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 217 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FEATURE:
 NAME/KEY: Peptide
 LOCATION: 32..53
 FEATURE:
 NAME/KEY: Peptide
 LOCATION: 94..115
 FEATURE:

NAME/KEY: Peptide
LOCATION: 133-153
FEATURE:
NAME/KEY: Peptide
LOCATION: 192-210
OTHER INFORMATION: /note: "the peptides above are
OTHER INFORMATION: depicted in Figure 1"

US-08-759-628-11

Query Match 57.8%: Score 461; DB 3; Length 217;
Best Local Similarity 70.3%: Pred. No. 4.4e-42;
Matches 102: Conservative 7; Mismatches 20; Indels 16; Gaps 4.

QY 2 FTPIPLSLFDNAMLRAHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNPOTLSFSSESPT 61
DB 27 FTPIPLSLFDNAMLRAHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNPOTLSFSSESPT 60
QY 62 PSNREETOOKSNLELRISLLIQSWLEPVLQIGTGFVFNCHLGS-----HLVE 111
DB 87 PSNREETOOKSNLELRISLLIQSWLEPVLQIGTGFVFNCHLGS-----HLVE 111
QY 112 ALYLVCG--ERGFFYTPKRGIVEQ 134
DB 146 GIOTLMGRLEDG---SPRTGQIFKQ 167

RESULT 9

US-09-284-878-1

Sequence 1: Application US/C92/4978
Patent No. 6342375
GENERAL INFORMATION:
APPLICANT: Olazaran, Martha Guetiero
APPLICANT: Saldaña, Hugo Barrera
APPLICANT: Salvador, Jose Maria Viader
TITLE OF INVENTION: Genetically Modified Melanotrophic β -galactosidase from the
TITLE OF INVENTION: Production and Secretion of the Human Growth Hormone
FILE REFERENCE: 1829.5010000
CURRENT FILING DATE: 1999-07-21
PRIOR APPLICATION NUMBER: US/09/284,878
PRIOR FILING DATE: 1997-10-24
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 1

TYPE: PRT

ORGANISM: Homo sapiens

US-09-284-878-1

Query Match 57.8%: Score 461; DB 4; Length 217;
Best Local Similarity 70.3%: Pred. No. 4.4e-42;
Matches 102: Conservative 7; Mismatches 20; Indels 16; Gaps 4.

QY 2 FTPIPLSLFDNAMLRAHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNPOTLSFSSESPT 61
DB 27 FTPIPLSLFDNAMLRAHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNPOTLSFSSESPT 60
QY 62 PSNREETOOKSNLELRISLLIQSWLEPVLQIGTGFVFNCHLGS-----HLVE 111
DB 87 PSNREETOOKSNLELRISLLIQSWLEPVLQIGTGFVFNCHLGS-----HLVE 111
QY 112 ALYLVCG--ERGFFYTPKRGIVEQ 134
DB 146 GIOTLMGRLEDG---SPRTGQIFKQ 167

RESULT 10

US-09-424-620B-25

Sequence 25: Application US/C94/2420B
Patent No. 6391585
GENERAL INFORMATION:
APPLICANT: HANIL SYNTHETIC FIBER CO., LTD.
JANG, KI-HYONG

NAME: Jae-Woong
RAE, Cheon-Soon
YANG, Doo-Suk
LEE, Jee-Won
SEONG, Baik-Lin
TITLE OF INVENTION: Process for preparing recombinant proteins using high-
efficiency expression vector from Saccharomyces cerevisiae.
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: BACHMAN & LAPORTE, P.C.
STREET: Suite 1201, 900 Chapel Street
City: New Haven
STATE: Connecticut
COUNTRY: U.S.A.
ZIP: 06510-2802
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
COMPUTER: IBM
OPERATING SYSTEM: WINDOWS 95/98
SOFTWARE: MS WORD
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/424,620B
FILING DATE: 24-Nov. 6391585-1999
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 241 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: PROTEIN
SEQUENCE DESCRIPTION: SFO ID NO: 25:
US-09-424-620B-25

Query Match 57.8%: Score 461; DB 4; Length 241;

Best Local Similarity 70.3%: Pred. No. 5e-42;

Matches 102: Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 2 FTPIPLSLFDNAMLRAHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNPOTLSFSSESPT 61

DB 51 FTPIPLSLFDNAMLRAHRLHQLAFDTYQFEFEAYIPKEQKYSFLQNPOTLSFSSESPT 110

QY 62 PSNREETOOKSNLELRISLLIQSWLEPVLQIGTGFVFNCHLGS-----HLVE 111

DB 111 PSNREETOOKSNLELRISLLIQSWLEPVLQIGTGFVFNCHLGS-----HLVE 111

QY 112 ALYLVCG--ERGFFYTPKRGIVEQ 134

DB 170 GIOTLMGRLEDG---SPRTGQIFKQ 191

RESULT 11

US-09-280-030-66

Sequence 66: Application US/09280030A

Patent No. 6506595

GENERAL INFORMATION:

APPLICANT: Sato, Seiji

APPLICANT: Higashikuni, Naohiko

APPLICANT: Kudo, Toshiyuki

APPLICANT: Kondo, Masazaki

TITLE OF INVENTION: DNAS ENCODING NEW FUSION PROTEINS AND PROCESSES FOR

TITLE OF INVENTION: PREPARING USEFUL POLYPEPTIDES THROUGH EXPRESSION OF THE

FILE REFERENCE: 382.1026

CURRENT APPLICATION NUMBER: US/09/280,030A

CURRENT FILING DATE: 1999-03-26

EARLIER APPLICATION NUMBER: JP10-87339/1998

EARLIER FILING DATE: 1998-03-31

NUMBER OF SEQ ID NOS: 66

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 66

LENGTH: 245

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Designated as
 OTHER INFORMATION: an amino acid sequence of WWPSP-WWPSP-23-TEV-Q-CH
 US-09-280-030-66

Query Match 57.8%; Score 461; DB 4; Length 245;
 Best Local Similarity 70.3%; Pred. No. 5.2e-42;
 Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 6;
 QY 2 FPIPLSLKFDNMLRAHRLHQAFTYQEFPPAYIPKEQKYSFLONPOTSLSPSESPT 61
 DE 55 FPIPLSLKFDNMLRAHRLHQAFTYQEFPPAYIPKEQKYSFLONPOTSLSPSESPT 114
 QY 62 PSNREETQKSNLELRISLLIQSWLEVPVQLGIGRFFVNHLOGS-----HLVE 111
 DE 115 PSNREETQKSNLELRISLLIQSWLEVPVQLGIGRFFVNHLOGS-----HLVE 111
 QY 112 AYLVCG--ERGFYTPKTRGIVEQ 134
 DE 174 GTQLKGRLEDC---SPRTGQIFKQ 195

RESULT 12

US-08-784-582-71
 Sequence 71, Application US/06794582
 Patent No. 6110707
 GENERAL INFORMATION:
 APPLICANT: Newgard, Christopher B.
 APPLICANT: Halban, Philippe A.
 APPLICANT: No. 6110707Mington, Karl D.
 APPLICANT: Clark, Samuel A.
 APPLICANT: Thiessen, Alice E.
 APPLICANT: Quade, Christian
 APPLICANT: Kruse, Fred
 APPLICANT: McGarry, Dennis
 TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
 NUMBER OF SEQUENCES: 79
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Arnold, White & Durkee
 STREET: P.O. Box 4433
 CITY: Houston
 STATE: Texas
 COUNTRY: USA
 ZIP: 77210
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/784,582
 FILING DATE: Concurrently Herewith
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/028,427
 FILING DATE: 15-OCT-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 09/589,028
 FILING DATE: 19-JAN-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Highlander, Steven L.
 REGISTRATION NUMBER: 37,642
 REFERENCE/DOCKET NUMBER: US/08/784,582
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 512/418-3000
 TELEFAX: 512/474-7577
 INFORMATION FOR SEQ ID NO: 71:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 274 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 US-08-784-582-71

Query Match 57.8%; Score 461; DB 3; Length 274;
 Best Local Similarity 70.3%; Pred. No. 6e-42;
 Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;
 QY 2 FPIPLSLKFDNMLRAHRLHQAFTYQEFPPAYIPKEQKYSFLONPOTSLSPSESPT 61
 DE 27 FPIPLSLKFDNMLRAHRLHQAFTYQEFPPAYIPKEQKYSFLONPOTSLSPSESPT 86
 QY 62 PSNREETQKSNLELRISLLIQSWLEVPVQLGIGRFFVNHLOGS-----HLVE 111
 DE 87 PSNREETQKSNLELRISLLIQSWLEVPVQLGIGRFFVNHLOGS-----HLVE 111
 QY 112 AYLVCG--ERGFYTPKTRGIVEQ 134
 DE 146 GTQLKGRLEDC---SPRTGQIFKQ 167

RESULT 13

US-08-784-582-73
 Sequence 73, Application US/08784582
 Patent No. 6110707
 GENERAL INFORMATION:
 APPLICANT: Newgard, Christopher B.
 APPLICANT: Halban, Philippe A.
 APPLICANT: No. 6110707Mington, Karl D.
 APPLICANT: Clark, Samuel A.
 APPLICANT: Thiessen, Alice E.
 APPLICANT: Quade, Christian
 APPLICANT: Kruse, Fred
 APPLICANT: McGarry, Dennis
 TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
 NUMBER OF SEQUENCES: 79
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Arnold, White & Durkee
 STREET: P.O. Box 4433
 CITY: Houston
 STATE: Texas
 COUNTRY: USA
 ZIP: 77210
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/784,582
 FILING DATE: Concurrently Herewith
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/028,427
 FILING DATE: 15-OCT-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/589,028
 FILING DATE: 19-JAN-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Highlander, Steven L.
 REGISTRATION NUMBER: 37,642
 REFERENCE/DOCKET NUMBER: US/08/784,582
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 512/418-3000
 TELEFAX: 512/474-7577
 INFORMATION FOR SEQ ID NO: 73:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 360 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 US-08-784-582-73

Query Match 57.8%; Score 461; DB 3; Length 360;
 Best Local Similarity 70.3%; Pred. No. 8.5e-42;

Matches 102: Conservative 7: Mismatches 20: Indels 11: Gaps 4:

QY 2 FPTTPSLRLEFNAMLRHRLHQLAFDTYQEFEEAYTPKEOKYSFIONPOTSLSFSESIFT 61
 : 111 1111 1111 1111 1111 111 111 111 111 111 111
 DB 27 FPTTPSLRLEFNAMLRHRLHQLAFDTYQEFEEAYTPKEOKYSFIONPOTSLSFSESIFT 61
 : 111 1111 1111 1111 1111 111 111 111 111 111 111

QY 62 PSNREETQOKSNLELLRISILLIQSMLEPVOLGTGPRFVNGHLCGS-----HLVE 111
 : 111 1111 1111 1111 1111 111 111 111 111 111 111
 DB 87 PSNREETQOKSNLELLRISILLIQSMLEPVOLGTGPRFVNGHLCGS-----HLVE 111
 : 111 1111 1111 1111 1111 111 111 111 111 111 111

QY 112 ALYLVCG--ERGFYTPKTRGIVEQ 134
 : 111 111 111 111 111 111 111 111 111 111 111
 DB 146 GIOTLMGRLEDG---SPRTGQIFKQ 167

RESULT 14
 US-09-465-461-1:
 : Sequence 1, Application US/09465461
 : Patent No. 634644
 : GENERAL INFORMATION:
 : APPLICANT: CHAPPEL, Scott
 : TITLE OF INVENTION: Human Growth Hormone to stimulate hematopoiesis and immune reorg
 : FILE REFERENCE: after hematopoietic stem cell transplantation in humans
 : CURRENT APPLICATION NUMBER: US/09/465,461
 : PRIOR FILING DATE: 1999-12-17
 : PRIOR APPLICATION NUMBER: 63/112,668
 : NUMBER OF SEQ ID NOS: 1
 : SOFTWARE: Patent in version 3.1
 : SEQ ID NO 1
 : LENGTH: 191
 : TYPE: PRT
 : ORGANISM: homo sapiens
 US-09-465-461-1

Query Match 57.3% Score 455: DB 1: Length 191:
 Best Local Similarity 69.7% Pred No. 2e-41:
 Matches 101: Conservative 7: Mismatches 21: Indels 18: Gaps 1:

QY 2 FPTTPSLRLEFNAMLRHRLHQLAFDTYQEFEEAYTPKEOKYSFIONPOTSLSFSESIFT 61
 : 111 1111 1111 1111 1111 111 111 111 111 111 111
 DB 1 FPTTPSLRLEFNAMLRHRLHQLAFDTYQEFEEAYTPKEOKYSFIONPOTSLSFSESIFT 61
 : 111 1111 1111 1111 1111 111 111 111 111 111 111

QY 62 PSNREETQOKSNLELLRISILLIQSMLEPVOLGTGPRFVNGHLCGS-----HLVE 111
 : 111 1111 1111 1111 1111 111 111 111 111 111 111
 DB 61 PSNREETQOKSNLELLRISILLIQSMLEPVOLGTGPRFVNGHLCGS-----HLVE 111
 : 111 1111 1111 1111 1111 111 111 111 111 111 111

QY 112 ALYLVCG--ERGFYTPKTRGIVEQ 134
 : 111 111 111 111 111 111 111 111 111 111 111
 DB 120 GIOTLMGRLEDG---SPRTGQIFKQ 167

RESULT 15
 US-09-187-756C-4
 : Sequence 4, Application US/08187756C
 : Patent No. 5597769
 : GENERAL INFORMATION:
 : APPLICANT: ROSEN, ET AL.
 : TITLE OF INVENTION: Human Growth Hormone
 : NUMBER OF SEQUENCES: 7
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: CAHELIA, BYRN, BAYN, SULFILLAN,
 : ADDRESS: 6 BECKER FARM ROAD
 : CITY: ROSELAND
 : STATE: NEW JERSEY
 : COUNTRY: USA
 : ZIP: 07068

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 INCH DISKETTE
 COMPUTER: IBM PS/2
 OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/187,756C
 FILING DATE: January 27, 1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: FERRARO, GREGORY D.
 REGISTRATION NUMBER: 36,134
 REFERENCE/DOCKET NUMBER: 325860-55
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201-994-1750
 TELEFAX: 201-994-1744
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 217 AMINO ACIDS
 TYPE: AMINO ACID
 STRANDEDNESS:
 TOPOLOGY: LINEAR
 MOLECULE TYPE: PROTEIN
 US-08-187-756C-4

Query Match 57.3% Score 455: DB 1: Length 217:
 Best Local Similarity 69.7% Pred No. 2e-41:
 Matches 101: Conservative 7: Mismatches 21: Indels 18: Gaps 4:

QY 2 FPTTPSLRLEFNAMLRHRLHQLAFDTYQEFEEAYTPKEOKYSFIONPOTSLSFSESIFT 61
 : 111 1111 1111 1111 1111 111 111 111 111 111 111
 DB 27 FPTTPSLRLEFNAMLRHRLHQLAFDTYQEFEEAYTPKEOKYSFIONPOTSLSFSESIFT 61
 : 111 1111 1111 1111 1111 111 111 111 111 111 111

QY 62 PSNREETQOKSNLELLRISILLIQSMLEPVOLGTGPRFVNGHLCGS-----HLVE 111
 : 111 1111 1111 1111 1111 111 111 111 111 111 111
 DB 87 PSNREETQOKSNLELLRISILLIQSMLEPVOLGTGPRFVNGHLCGS-----HLVE 111
 : 111 1111 1111 1111 1111 111 111 111 111 111 111

QY 112 ALYLVCG--ERGFYTPKTRGIVEQ 134
 : 111 111 111 111 111 111 111 111 111 111 111
 DB 146 GIOTLMGRLEDG---SPRTGQIFKQ 167

Search completed: September 16, 2003, 12:41:25
 CPU time: 24.93 secs

GenCore version 5.1.1.6
 Copyright (c) 1993 - 2003, Compugen Ltd.

OM protein - protein search, using sw mode.

Run on: September 16, 2003, 12:38:55 : Search time 47.2753 seconds
 (without alignments)
 472.415 Million cell updates/sec

Title: US-09-423-100-7

Perfect score: 797

Sequence: 1 MFPTPLSKIFDNYLRAER.....IVGQCTSTCSLYLHENVCK 150

Scoring table: BLSUM62

Gapop 10.0 , Gapext 0.5

Searched: 556259 seqs, 148393369 residues

Total number of hits satisfying chosen parameters: 556265

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMP.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*

Pred. No. is the number of residues predicted by Clustal to have a
 score greater than or equal to the score of the result being printed.
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	797	100.0	150	14	US-10-054-873-7
2	555.5	69.7	107	14	US-10-054-873-6
3	470	59.0	92	14	US-10-054-873-2
4	461	57.8	191	11	US-09-984-010-23
5	461	57.8	191	12	US-10-153-207-1
6	461	57.8	191	12	US-10-403-377-1
7	461	57.8	191	12	US-10-456-708-1
8	461	57.8	191	12	US-10-238-148-1
9	461	57.8	214	12	US-10-153-207-6
10	461	57.8	217	9	US-09-929-918-9
11	461	57.6	245	9	US-09-280-035-66
12	455	57.1	217	9	US-09-853-689-2
13	455	57.1	217	12	US-09-969-7482-4
14	454	57.0	217	10	US-09-804-409A-16
15	448	56.2	217	5	US-09-853-688-4

10	445	55.4	291	12	US-09-824-200-12
17	440	55.2	191	12	US-10-300-822-1
18	433	54.3	191	12	US-10-289-845-3
19	399	50.1	217	9	US-09-850-887-3
20	381	47.8	163	15	US-10-043-487-350
21	381	47.8	191	12	US-10-153-207-2
22	373	46.4	249	15	US-10-103-913-411
23	345	43.3	246	15	US-10-188-246-18
24	301.5	37.8	190	12	US-10-153-207-3
25	294.5	37.0	138	10	US-09-861-687-19
26	294	36.9	52	14	US-10-054-873-5
27	286	35.9	147	9	US-09-736-611-8
28	286	35.9	147	9	US-09-740-359-7
29	284.5	35.7	124	10	US-09-894-711-18
30	284.5	35.7	144	9	US-09-736-611-6
31	284.5	35.7	144	9	US-09-740-359-5
32	284.5	35.7	146	10	US-09-894-711-5
33	278.5	34.9	51	11	US-09-858-935B-5
34	278.5	34.9	51	14	US-10-028-410-3
35	278	34.9	117	9	US-09-280-030-63
36	277	34.8	96	10	US-09-947-563-4
37	275.5	34.6	124	9	US-09-736-611-12
38	275.5	34.6	124	9	US-09-740-359-12
39	275.5	34.6	124	10	US-09-894-711-12
40	275.5	34.6	125	9	US-09-736-611-10
41	275.5	34.6	125	9	US-09-740-359-10
42	275.5	34.6	125	10	US-09-894-711-10
43	273	34.5	50	14	US-10-056-009A-3
44	271	34.0	96	10	US-09-947-563-5
45	270	33.9	110	9	US-09-205-658-125

ALIGNMENTS

RESULT 1
 US-10-054-873-7
 : Sequence 7, Application US/10054873
 : Publication No. US20020164712A1
 : GENERAL INFORMATION:
 : APPLICANT: Can. Zhong Hu
 : TITLE OF INVENTION: Chimeric Protein Containing an
 : Intramolecular Chaperone-Like Sequence
 : NUMBER OF SEQUENCES: 7
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Townsend and Townsend and Crew LLP
 : STREET: Two Embarcadero Center, Eighth Floor
 : CITY: San Francisco
 : STATE: California
 : COUNTRY: USA
 : ZIP: 94111-3834
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/0/054.873
 : FILING DATE: 22-Jan-2002
 : CLASSIFICATION: <Unknown>
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: WO PCT/CN98/00052
 : FILING DATE: 31-MAR-1998
 : APPLICATION NUMBER: US 09/423,100
 : FILING DATE: 11-DEC-2000
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Mycroft, Frank J
 : REGISTRATION NUMBER: 45,946
 : REFERENCE/DOCKET NUMBER: 020167-000130US
 : INFORMATION FOR SEQ ID NO: 7:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 150 amino acids
 : TYPE: amino acid

Sequence 12, Appl
 Sequence 1, Appl
 Sequence 3, Appl
 Sequence 3, Appl
 Sequence 350, Appl
 Sequence 2, Appl
 Sequence 411, Appl
 Sequence 18, Appl
 Sequence 3, Appl
 Sequence 19, Appl
 Sequence 5, Appl
 Sequence 8, Appl
 Sequence 7, Appl
 Sequence 18, Appl
 Sequence 6, Appl
 Sequence 5, Appl
 Sequence 5, Appl
 Sequence 5, Appl
 Sequence 3, Appl
 Sequence 4, Appl
 Sequence 12, Appl
 Sequence 12, Appl
 Sequence 10, Appl
 Sequence 10, Appl
 Sequence 10, Appl
 Sequence 5, Appl
 Sequence 5, Appl
 Sequence 125, App

STRANDEDNESS: <Unknown>
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 7:
 US-10-054-873-7

Query Match: 100.0%; Score 797; LP 14; Length 150;
 Best Local Similarity 100.0%; Pred. No. 1.4e-40;
 Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 MFPTPLSRFDNAMLRAHRLHQLAFDTYQEFAYIPKEQKYSFLQNP 60
 DB 1 MFPTPLSRFDNAMLRAHRLHQLAFDTYQEFAYIPKEQKYSFLQNP 60
 QY 61 TPSNREETQOKSNLELRISLLIOSKLEPVOLGTGSPFVNOHLCGSHLVEALYLVCGER 120
 DB 61 TPSNREETQOKSNLELRISLLIOSKLEPVOLGTGSPFVNOHLCGSHLVEALYLVCGER 120
 QY 121 GFYTPKTRGIVEQCCISCSLYQENYCN 150
 DB 121 GFYTPKTRGIVEQCCISCSLYQENYCN 150

RESULT 2

US-10-054-873-6

Sequence 6, Application US/10054873
 Publication No. US20020164712A1

GENERAL INFORMATION:

APPLICANT: Gan, Zhong R
 TITLE OF INVENTION: Chimeric Protein Containing an Intramolecular Chaperone-Like Sequence

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP
 STREET: Two Embarcadero Center, Eighth Floor
 CITY: San Francisco
 STATE: California
 COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: P-OS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/054,873

FILING DATE: 22-Jan-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/CN98/00052

FILING DATE: 31-MAR-1998

APPLICATION NUMBER: US 09/423,100

FILING DATE: 11-DEC-2000

ATTORNEY/AGENT INFORMATION:

NAME: Mycroft, Frank J

REGISTRATION NUMBER: 46,946

REFERENCE/DOCKET NUMBER: 020167-00013005

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 107 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-10-054-873-6

Query Match:

Best Local Similarity 71.3%; Score 555.5; DB 14; Length 107;
 Matches 107; Conservative 0; Mismatches 0; Indels 4; Gaps 1

QY 1 MFPTPLSRFDNAMLRAHRLHQLAFDTYQEFAYIPKEQKYSFLQNP 60
 DB 1 MFPTPLSRFDNAMLRAHRLHQLAFDTYQEFAYIPKEQKYSFLQNP 60

DB 1 MFPTPLSRFDNAMLRAHRLHQLAFDTYQEFAYIPKEQKYSFLQNP 49
 QY 61 TPSNREETQOKSNLELRISLLIOSKLEPVOLGTGSPFVNOHLCGSHLVEALYLVCGER 120
 DB 61 TPSNREETQOKSNLELRISLLIOSKLEPVOLGTGSPFVNOHLCGSHLVEALYLVCGER 120
 QY 121 GFYTPKTRGIVEQCCISCSLYQENYCN 150
 DB 78 GFYTPKTRGIVEQCCISCSLYQENYCN 107

RESULT 3

US-10-054-873-2

Sequence 2, Application US/10054873

Publication No. US20020164712A1

GENERAL INFORMATION:

APPLICANT: Gan, Zhong R

TITLE OF INVENTION: Chimeric Protein Containing an Intramolecular Chaperone-Like Sequence

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP
 STREET: Two Embarcadero Center, Eighth Floor
 CITY: San Francisco
 STATE: California
 COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: P-OS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/054,873

FILING DATE: 22-Jan-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/CN98/00052

FILING DATE: 31-MAR-1998

APPLICATION NUMBER: US 09/423,100

FILING DATE: 11-DEC-2000

ATTORNEY/AGENT INFORMATION:

NAME: Mycroft, Frank J

REGISTRATION NUMBER: 46,946

REFERENCE/DOCKET NUMBER: 020167-00013005

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 92 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-10-054-873-2

Query Match: 59.0%; Score 470; DB 14; Length 92;

Best Local Similarity 100.0%; Pred. No. 1.5e-44;

Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFPTPLSRFDNAMLRAHRLHQLAFDTYQEFAYIPKEQKYSFLQNP 60
 DB 1 MFPTPLSRFDNAMLRAHRLHQLAFDTYQEFAYIPKEQKYSFLQNP 60
 QY 61 TPSNREETQOKSNLELRISLLIOSKLEPVQ 92
 DB 61 TPSNREETQOKSNLELRISLLIOSKLEPVQ 92

RESULT 4

US-09-964-010-23

Sequence 23, Application US/09984010

Publication No. US20030104578A1

GENERAL INFORMATION:

```

APPLICANT: BALANCE, DAVID JAMES
TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GRAFT HEPATITIS B AND SERUM ALBUMIN
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: PINKHAM, HENDERSON, PARAHOW, GARRETT & KENNEDY, LLP
STREET: 1306 I Street, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent's Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/884,010
FILING DATE: 21-May-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 05/092,873
FILING DATE: 25-JUN-1998
APPLICATION NUMBER: PCT/SB96/03164
FILING DATE: 19-DEC-1995
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
HYDROTHERICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-584-913-23

Query Match      57.8%   Score 461    BR 12    Length 191
Best Local Similarity 70.3%   Pred. No. 3,93-43
Matches 102     Conservative 7    Mismatches 20    Indels 143 gaps 4

QY  2 FTPLSLLEFNAMLAHRLHQALFDYVQHFEENAIPEKQKYSLONPQTSSISEPSIPT 61
DB  1 DIIIIIIIIIII IIIIIIIIIIII IIII II IIII IIII IIII IIII
QY  62 PSNRRTCKSKNLSLRLRSLILGCSMLRPGLTGTPRFVNCHAGS-----ELVF 111
DB  61 ESNRRETCKSKNLSLRLRSLLIGCSMLRPGLTSLSPAFNSIVGTASTNVNLKKLR 111
QY  112 AYLVEG--ERRFFVTKTRGIIPG 134
DB  120 GIGTIMGGIEFG---SPITGSIIRKG 141

RESULT 5
US-10-153-207-1
Sequence 1, Application: US/10153207
Publication No.: US20030153093A1
GENERAL INFORMATION:
APPLICANT: James A. Wells
APPLICATOR: Brian C. Cunningham
TITLE OF INVENTION: GROWTH HORMONE VARIANTS
FILE REFERENCE: 569, 12-US-C7
CURRENT APPLICATION NUMBER: US/10/-53,207
CURRENT FILING DATE: 2002-05-22
PRIOR APPLICATION NUMBER: 08/479,884
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 08/190,723
PRIOR FILING DATE: 1994-02-02
PRIOR APPLICATION NUMBER: 07/960,227
PRIOR FILING DATE: 1992-10-13
PRIOR APPLICATION NUMBER: 07/875,204
PRIOR FILING DATE: 1992-04-27
PRIOR APPLICATION NUMBER: 07/428,066
```



```

1  APPLICANT: Cox III, George N
2  TITLE OF INVENTION: Derivatives of Growth Hormone and Related Proteins
3  FILE REFERENCE: 4152-1-PUS
4  CURRENT APPLICATION NUMBER: US/12/400,728
5  CURRENT FILING DATE: 2003-03-26
6  PRIOR APPLICATION NUMBER: US/09/462,941
7  PRIOR FILING DATE: 2002-01-14
8  PRIOR APPLICATION NUMBER: 60/052,516
9  PRIOR FILING DATE: 1997-07-14
10 NUMBER OF SEQ ID NOS: 41
11 SOFTWARE: PatentIn Ver. 2.0
12 SEQ ID NO 1
13 LENGTH: 191
14 TYPE: PRT
15 ORGANISM: Homo sapiens
16 US-10-400-708-1

Query Match      57.8%  Score 461  DB 12:  Length 191
Best local Similarity 70.3%  Pred. No. 3.9e-43
Matches 102:  Conservative  7;  Mismatches  20;  Indels  16;  Gaps  4

QY  2  FPTPLSLRFDNMLRAHCHQIAEDYQEFEEAYIPKEQKYSFQNPOTSLSPESIPT 61
DB  1  FPTPLSLRFDNMLRAHCHQIAEDYQEFEEAYIPKEQKYSFQNPOTSLSPESIPT 60
QY  62  PSNREETQOKSNLELRISLLIQSWLEPVQGTGPRFVNOHLCSS-----HLVE 111
DB  61  PSNREETQOKSNLELRISLLIQSWLEPVQGTGPRFVNOHLCSS-----HLVE 110
QY  112  ALYLVG--ERGFYTPKTRGIVEQ 134
DB  120  GQTLNGRLDG---SPRTGQIFKO 141

```

```

RESULT 8
US-10-296-148-1
1  Sequence 1, Application US/16226148
2  Publication No. US20030171284A1
3  GENERAL INFORMATION:
4  APPLICANT: Holder Biotechnology, Inc.
5  TITLE OF INVENTION: Derivatives of Growth Hormone and Related Proteins
6  FILE REFERENCE: 4152-1-PUS
7  CURRENT APPLICATION NUMBER: US/10/298,148
8  CURRENT FILING DATE: 2002-11-15
9  PRIOR APPLICATION NUMBER: US/09/462,941
10 PRIOR FILING DATE: 2003-01-14
11 PRIOR APPLICATION NUMBER: 60/052,516
12 PRIOR FILING DATE: 1997-07-14
13 NUMBER OF SEQ ID NOS: 41
14 SOFTWARE: PatentIn Ver. 2.0
15 SEQ ID NO 1
16 LENGTH: 191
17 TYPE: PRT
18 ORGANISM: Homo sapiens
19 US-10-298-148-1

Query Match      57.6%  Score 461  DB 12:  Length 191
Best local Similarity 70.3%  Pred. No. 3.9e-43
Matches 102:  Conservative  7;  Mismatches  20;  Indels  16;  Gaps  4

QY  2  FPTPLSLRFDNMLRAHCHQIAEDYQEFEEAYIPKEQKYSFQNPOTSLSPESIPT 61
DB  1  FPTPLSLRFDNMLRAHCHQIAEDYQEFEEAYIPKEQKYSFQNPOTSLSPESIPT 60
QY  62  PSNREETQOKSNLELRISLLIQSWLEPVQGTGPRFVNOHLCSS-----HLVE 111
DB  61  PSNREETQOKSNLELRISLLIQSWLEPVQGTGPRFVNOHLCSS-----HLVE 110
QY  112  ALYLVG--ERGFYTPKTRGIVEQ 134
DB  120  GQTLNGRLDG---SPRTGQIFKO 141

```

```

RESULT 9
US-10-153-207-5
1  Sequence 6, Application US/10153207
2  Publication No. US20030153003A1
3  GENERAL INFORMATION:
4  APPLICANT: James A. Wells
5  APPLICANT: Brian C. Cunningham
6  TITLE OF INVENTION: GROWTH HORMONE VARIANTS
7  FILE REFERENCE: 569-12-US-C7
8  CURRENT APPLICATION NUMBER: US/10/153,207
9  CURRENT FILING DATE: 2002-05-22
10 PRIOR APPLICATION NUMBER: 08/479,884
11 PRIOR FILING DATE: 1995-06-07
12 PRIOR APPLICATION NUMBER: 06/190,723
13 PRIOR FILING DATE: 1994-02-02
14 PRIOR APPLICATION NUMBER: 07/966,227
15 PRIOR FILING DATE: 1992-10-13
16 PRIOR APPLICATION NUMBER: 07/875,204
17 PRIOR FILING DATE: 1992-04-27
18 PRIOR APPLICATION NUMBER: 07/428,066
19 PRIOR FILING DATE: 1989-10-26
20 PRIOR APPLICATION NUMBER: 07/264,611
21 PRIOR FILING DATE: 1988-10-28
22 NUMBER OF SEQ ID NOS: 20
23 SOFTWARE: FastSeq for Windows Version 4.0
24 SEQ ID NO 6
25 LENGTH: 214
26 TYPE: PRT
27 ORGANISM: Homo Sapiens
28 US-10-153-207-5

Query Match      57.8%  Score 461  DB 12:  Length 214
Best local Similarity 70.3%  Pred. No. 4.5e-43
Matches 102:  Conservative  7;  Mismatches  20;  Indels  16;  Gaps  4

QY  2  FPTPLSLRFDNMLRAHCHQIAEDYQEFEEAYIPKEQKYSFQNPOTSLSPESIPT 61
DB  24  FPTPLSLRFDNMLRAHCHQIAEDYQEFEEAYIPKEQKYSFQNPOTSLSPESIPT 83
QY  62  PSNREETQOKSNLELRISLLIQSWLEPVQGTGPRFVNOHLCSS-----HLVE 111
DB  84  PSNREETQOKSNLELRISLLIQSWLEPVQGTGPRFVNOHLCSS-----HLVE 142
QY  112  ALYLVG--ERGFYTPKTRGIVEQ 134
DB  143  GQTLNGRLDG---SPRTGQIFKO 164

```

```

RESULT 10
US-09-929-918-9
1  Sequence 9, Application US/09929918
2  Patent No. US20020090678A1
3  GENERAL INFORMATION:
4  APPLICANT: Kordyum, Vitaliy A.
5  APPLICANT: Cherdnykh, Svetlana I.
6  APPLICANT: Slavchenko, Irina Yu.
7  APPLICANT: Vozianov, Aleksandr
8  TITLE OF INVENTION: PHAGE-DEPENDENT SUPER PRODUCTION OF
9  BIOLOGICALLY ACTIVE PROTEIN AND PEPTIDES
10 FILE REFERENCE: PHAGE.006A
11 CURRENT APPLICATION NUMBER: US/09/929,918
12 CURRENT FILING DATE: 2001-08-15
13 PRIOR APPLICATION NUMBER: 09/318,288
14 PRIOR FILING DATE: 1999-05-25
15 NUMBER OF SEQ ID NOS: 11
16 SOFTWARE: FastSeq for Windows Version 4.0
17 SEQ ID NO 9
18 LENGTH: 217
19 TYPE: PRT
20 ORGANISM: Homo sapiens
21 US-09-929-918-9

```


Gene version 5.1.6
Copyright (c) 1993 - 2003 Computer, Ltd.

OM protein - protein search, using sw mode:

Run on: September 16, 2003, 12:36:15 : Search time 17.5957 seconds
(without alignments)
823,845 Million cell updates/sec

Title: US-09-423-100-7

Perfect score: 797

Sequence: 1 MFPTPLSRLEFDMALRAHR.....IVECCISICSLACLENEN 150

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 95158652 residues

Total number of hits satisfying chosen parameters: 283305

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 43 summaries

Database : PIR_76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	461	57.8	217	1	STHUC
2	460	57.7	217	2	somatotropin 1, pre
3	426.5	53.5	217	1	STHUV
4	426.5	53.5	256	1	somatotropin 2, pre
5	407.5	51.1	217	2	STHUC
6	405	50.8	217	2	STHUC
7	396	49.7	217	2	STHUC
8	396	49.7	217	2	STHUC
9	381	47.8	217	1	STHUC
10	381	47.8	217	2	STHUC
11	359.5	45.1	215	2	STHUC
12	310.5	39.0	216	2	STHUC
13	307.5	38.6	190	2	STHUC
14	306.5	38.5	190	1	STHUC
15	304.5	38.2	216	1	STHUC
16	302.5	38.0	216	1	STHUC
17	302.5	38.0	216	2	STHUC
18	301.5	37.8	190	2	STHUC
19	301.5	37.8	216	1	STHUC
20	301.5	37.8	216	2	STHUC
21	301.5	37.8	216	2	STHUC
22	299.5	37.6	216	2	STHUC
23	297.5	37.3	190	1	STHUC
24	295.5	37.1	190	2	STHUC
25	289.5	36.3	217	1	STHUC
26	289.5	36.3	217	1	STHUC
27	289.5	36.3	217	1	STHUC
28	289.5	36.3	217	2	STHUC
29	278.5	34.9	216	2	STHUC

30 277.5 34.8 110 1 INRB
31 277.5 34.8 110 2 B42179
32 275.5 34.6 226 2 A50509
33 275.5 34.5 98 2 PC7082
34 273.5 34.3 51 1 INWRP
35 273.5 34.3 51 1 INWRP
36 273.5 34.3 51 1 INEC
37 273.5 34.3 110 2 J00176
38 272.5 34.1 110 2 A42179
39 271.5 34.1 51 1 INHY
40 270 33.9 110 1 IFHU
41 268.5 33.7 51 1 INMSSP
42 268.5 33.7 131 2 A60625
43 267.5 33.6 51 2 A59151
44 266.5 33.4 105 1 IFBO
45 265.5 33.3 110 2 148166

ALIGNMENTS

RESULT 1

STHUC

somatotropin 1 precursor [validated] - human
N:Alternate names: growth hormone 1; hGH-N; pituitary somatotropin
N:Contains: growth hormone 5K peptide; somatotropin 1, long form; somatotropin 1, s
C:Species: Homo sapiens (man)
C:Date: 24-Apr-1984 #sequence_revision 10-Feb-1995 #text_change 08-Dec-2000
C:Accession: A93731; A32435; A93694; A94247; A90051; A93778; A91764; A90217
R:Accession: A93731; A32435; A93694; A94247; A90051; A93778; A91764; A90217
Nucleic Acids Res. 9, 3719-3730, 1981
A:Title: Human growth hormone DNA sequence and mRNA structure: possible alternative
A:Reference number: A93731; MUID:82014939; PMID:6269091
A:Accession: A93731
A:Molecule type: DNA
A:Residues: 1-217 <DEN>
A:Cross-references: GH:V00520
A:Note: The 20K short form somatotropin lacks residues 58-72 (32-46 in the active h
R:Chou, E.Y.; Liao, Y.C.; Smith, D.H.; Barrera-Saldana, H.A.; Gellinas, R.E.; Seebur
Genomics 4, 479-497, 1989
A:Title: The human growth hormone locus: nucleotide sequence, biology, and evolution
A:Reference number: A32435; MUID:89307277; PMID:2744760
A:Accession: A32435
A:Molecule type: DNA
A:Residues: 1-217 <CHE>
A:Cross-references: GH:V00519
R:Koska, W.; Rougeon, F.
Nucleic Acids Res. 7, 305-320, 1979
A:Title: Molecular cloning and nucleotide sequence of the human growth hormone struc
A:Reference number: A93694; MUID:80034477; PMID:385281
A:Accession: A93694
A:Molecule type: mRNA
A:Residues: 1-217 <ROS>
A:Cross-references: GH:V00519
A:Note: 35,950 was also found
R:Marlet, J.A.; Hallowell, R.A.; Baxter, J.D.; Goodman, H.M.
Science 205, 602-607, 1979
A:Title: Human growth hormone: complementary DNA cloning and expression in bacteria
A:Reference number: A94247; MUID:79203293; PMID:377496
A:Accession: A94247
A:Molecule type: mRNA
A:Residues: 1-217 <MAR>
R:Li, C.H.; Dixon, J.S.; Liu, W.K.
Arch. Biochem. Biophys. 133, 70-91, 1969
A:Title: Human pituitary growth hormone. XIX. The primary structure of the hormone.
A:Reference number: A90048; MUID:59289202; PMID:5810834
A:Contents: annotation
R:Li, C.H.; Dixon, J.S.
Arch. Biochem. Biophys. 146, 233-236, 1971
A:Title: Human pituitary growth hormone. XXII. The primary structure of the hormone
A:Reference number: A90051; MUID:72143935; PMID:5144027
A:Accession: A90051
A:Molecule type: protein

Query Match	57.7%	Score 460:	DB 2:	Length 217;
Best Local Similarity	98.9%;	Pred. No. 9.4e-38;		
Matches	90:	Conservative 0;	Mismatches 1;	Indels 0; Gaps 0;
QY	2	FPPIPSRIFDNLRAHLRHQLAFDYQGFEEAY:PKFKYSFLLQNPQTSLSFESIPT	61	
DB	27	FPPIPSRIFDNLRAHLRHQLAFDYQGFEEAY:PKFKYSFLLQNPQTSLSFESIPT	86	
QY	62	PSNRRETOQKSNLELLKISILLIQSWLEPVQ	92	

Db 87 PSNREETQOKSNLELLRISLLIQSWLEPVQL-----GTGPRFVNQRL 117

RESULT 3

SHUV

somatotropin 2 precursor - human

N:Alternate names: growth hormone variant-2; placental somatotropin form 2

C:Species: Homo sapiens (man)

C:Date: 17-Dec-1982 #sequence_revision 10-Feb-1995 #text_change 21-Jul-1995

C:Accession: D32433; M28072; A31511; 152163; A60711

R:Chen, F.Y.; Xiao, Y.C.; Smith, D.E.; Barrera Saldaña, H.A.; Gallinas, R. J. Genomics 4, 479-497, 1989

A:Title: The human growth hormone locus: nucleotide sequence, cloning, and evolution.

A:Reference number: A32435; MUID:8307277; PMID:2743768

A:Accession: D32433

A:Molecule type: DNA

A:Residues: 1-217 <CHE>

A:Cross-references: G3:J03071; NID:8183148; PIDN:AAA52552.1; PIDN:LEF152

R:Coake, N.E.; Ray, J.; Emery, J.G.; Liebhaber, S.A.

J. Biol. Chem. 263, 9001-9006, 1988

A:Title: Two distinct species of human growth hormone-variant mRNA in the human placenta

A:Reference number: A92725; MUID 86243769; PMID:3379057

A:Accession: B28072

A:Molecule type: mRNA

A:Residues: 1-217 <COO>

R:Seeburg, P.H.

DNA 1, 239-249, 1982

A:Title: The human growth hormone gene family: nucleotide sequences show recent divergence

A:Reference number: A61511; MUID:8182010; PMID:7165039

A:Accession: A61511

A:Molecule type: DNA

A:Residues: 1-347; 36-217 <SEP>

R:Iqbal, A.; Scippo, M.L.; Frankenne, F.; Hennan, G.

Arch. Int. Physiol. Biochim 96, 63-67, 1988

A:Title: Cloning and nucleotide sequence of placenta, hGH-V cDNA.

A:Reference number: 152104; MUID:89024984; PMID:2460050

A:Accession: 152104

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-217 <U>

A:Cross-references: GB:M84851; NID:8183148; PIDN:AAA52552.1; PIDN:LEF152

R:Frankenne, F.; Scippo, M.L.; Van Boesum, J.; Iqbal, A.; Hennan, G.

J. Clin. Endocrinol. Metab. 71, 1518, 1990

A:Title: Identification of placental human growth hormone as the at with hormone V gene.

A:Reference number: A60711; MUID:8317018; PMID:2166278

A:Accession: A60711

A:Molecule type: protein

A:Residues: 27-44; 46-57 <RA>

A:Experimental source: tissue placenta

A:Note: partial glycosylation was demonstrated by protein binding

C:Comment: This gene is expressed by the placenta.

C:Genetics:

A:Gene: GDB:G2

A:Cross-references: GDB:119983; OMIM:139240

A:Map position: 17q22-17q24

A:Introns: 4/1; 57/3; 97/3; 152/3

C:Superfamily: prolactin

C:Keywords: alternative splicing; glycoprotein; hormone; placenta

F:1-26/Domain: signal sequence #status predicted <SIG>

F:27-217/product: somatotropin 2, long splice form #status predicted <SIG>

F:27-57/3-217/product: somatotropin 2, short splice form #status predicted <SIG>

F:79-191/208-215/disulfide bonds: #status predicted

F:166/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 53.5% Score 426.5; DB 1; Length 217;
Best Local Similarity 78.1% Pred. No. 1.8e-34;
Matches 89; Conservative 4; Mismatches 10; Indels 11; Gaps 1;

QY 2 FPTPLSLFDNMLRAHRLHQLAFDTYCEFEAYIPKEOKYSFLONPOTSLSFSESIPT 61
|||||
Db 27 FPTPLSLFDNMLRAHRLHQLAFDTYCEFEAYIPKEOKYSFLONPOTSLSFSESIPT 61

QY 62 PSNREETQOKSNLELLRISLLIQSWLEPVQL-----GTGPRFVNQRL 104
|||||
Db 87 PSNREETQOKSNLELLRISLLIQSWLEPVQLRVSFVNSLVYGASDSNVYRHL 140

RESULT 4

SHUV

somatotropin 2 precursor, splice form 2 - human

N:Alternate names: growth hormone variant-2; placental somatotropin form 2

C:Species: Homo sapiens (man)

C:Date: 30-Sep-1989 #sequence_revision 10-Feb-1995 #text_change 02-Sep-1997

C:Accession: A28072

R:Coake, N.E.; Ray, J.; Emery, J.G.; Liebhaber, S.A.

J. Biol. Chem. 263, 9001-9006, 1988

A:Title: Two distinct species of human growth hormone-variant mRNA in the human placenta

A:Reference number: A92725; MUID:86243769; PMID:3379057

A:Accession: A28072

A:Molecule type: mRNA

A:Residues: 1-256 <COO>

A:Note: an alternative splice junction for intron 4 is used

C:Genetics:

A:Gene: GDB:G2

A:Cross-references: GDB:119983; OMIM:139240

A:Map position: 17q22-17q24

A:Introns: 4/1; 57/3; 97/3; 152/3

C:Superfamily: prolactin

C:Keywords: alternative splicing; hormone; placenta

F:1-26/Domain: signal sequence #status predicted <SIG>

F:27-256/product: somatotropin 2 splice form 2 #status predicted <NAT>

Query Match 53.5% Score 426.5; DB 1; Length 256;
Best Local Similarity 78.1% Pred. No. 2.2e-34;
Matches 89; Conservative 4; Mismatches 10; Indels 11; Gaps 1;

QY 2 FPTPLSLFDNMLRAHRLHQLAFDTYCEFEAYIPKEOKYSFLONPOTSLSFSESIPT 61
Db 27 FPTPLSLFDNMLRAHRLHQLAFDTYCEFEAYIPKEOKYSFLONPOTSLSFSESIPT 86

QY 62 PSNREETQOKSNLELLRISLLIQSWLEPVQL-----GTGPRFVNQRL 104
|||||
Db 87 PSNREETQOKSNLELLRISLLIQSWLEPVQLRVSFVNSLVYGASDSNVYRHL 140

RESULT 5

SHUV

chorionic somatotropin-3 - rhesus macaque

C:Species: Macaca mulatta (rhesus macaque)

C:Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 16-Jul-1999

C:Accession: 167409

R:Gelos, T.G.; Durning, M.; Fisher, J.M.; Fowler, P.D.

Endocrinology 133, 1744-1752, 1994

A:Title: Cloning of four growth hormone/chorionic somatomammotropin-related cDNAs

A:Reference number: 153267; MUID:94008724; PMID:8404617

A:Accession: 167409

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-217 <RES>

A:Cross-references: GDB:16554; NID:9293112; PIDN:AAA18841.1; PID:9293113

C:Superfamily: prolactin

Query Match 51.1% Score 407.5; DB 2; Length 217;
Best Local Similarity 71.9% Pred. No. 1.4e-32;
Matches 82; Conservative 12; Mismatches 19; Indels 1; Gaps 1;

QY 3 FTLPLSLFDNMLRAHRLHQLAFDTYCEFEAYIPKEOKYSFLONPOTSLSFSESIPT 62
|||||
Db 28 PSVPLSLRDLNMMHRLHQLAFDTYCEFEAYIPKEOKYSFLONPOTSLSFSESIPT 87

QY 63 SNREETQOKSNLELLRISLLIQSWLEPVQLRVSFVNSLVYGASDSNVYRHL 116
|||||
Db 88 SNREETQOKSNLELLRISLLIQSWLEPVQL-LGSVFANNLVYGTSESAYOLL 140

```

RESULT 6
I67411
somatotropin - rhesus macaque
N:Alternate names: growth hormone
C:Species: Macaca mulatta (rhesus macaque)
C>Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 16-Jul-1999
C:Accession: I67411
R:Golos, T.G.; Durning, M.; Fisher, J.M.; Fowler, P.D.
Endocrinology 133, 1744-1752, 1993
A:Title: Cloning of four growth hormone/chorionic somatomammotropin-related complementary DNAs from the placenta of rhesus macaques
A:Reference number: 153267; MID:94008724; PMID:8404617
A:Accession: I67411
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-217 <RES>
A:Cross-references: GB:L16553; NID:q293116; PID:AAA20180.1; PID:q293117
C:Superfamily: prolactin

Query Match 50.8%; Score 405; DB 2; Length 217;
Best Local Similarity 67.7%; Pred. No. 2.4e-32;
Matches 86; Conservative 9; Mismatches 18; Indels 14; Gaps 2;

QY 2 PPTPLSKLFDNAMI RAHRLHQLAFDTYQFEFEAYIPKQKYSFLQNPQTSLSSESISPT 62
DQ 27 PPTPLSKLFDNAMI RAHRLHQLAFDTYQFEFEAYIPKQKYSFLQNPQTSLSSESISPT 62
QY 62 PSNREETQOKSNLELRISLLIQLSQWLEPVQ 92
DQ 87 PSNREETQOKSNLELRISLLIQLSQWLEPVQ 92
QY 115 INCCGK 121
DQ 140 LKKLES 146

RESULT 7
I67408
chorionic somatomammotropin-2 - rhesus macaque (fragment)
C:Species: Macaca mulatta (rhesus macaque)
C>Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 16-Jul-1999
C:Accession: I67408
R:Golos, T.G.; Durning, M.; Fisher, J.M.; Fowler, P.D.
Endocrinology 133, 1744-1752, 1993
A:Title: Cloning of four growth hormone/chorionic somatomammotropin-related complementary DNAs from the placenta of rhesus macaques
A:Reference number: 153267; MID:94008724; PMID:8404617
A:Accession: I67408
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-212 <RES>
A:Cross-references: GB:L16553; NID:q293110; PID:AAA18849.1; PID:q293111
C:Superfamily: prolactin

Query Match 49.7%; Score 396; DB 2; Length 212;
Best Local Similarity 82.2%; Pred. No. 1.8e-31;
Matches 74; Conservative 11; Mismatches 5; Indels 0; Gaps 0;

QY 3 PPTPLSKLFDNAMI RAHRLHQLAFDTYQFEFEAYIPKQKYSFLQNPQTSLSSESISPT 62
DQ 23 PTPPLSKLFDNAMI RAHRLHQLAFDTYQFEFEAYIPKQKYSFLQNPQTSLSSESISPT 62
QY 63 SNREETQOKSNLELRISLLIQLSQWLEPVQ 92
DQ 83 SNREETQOKSNLELRISLLIQLSQWLEPVQ 112

RESULT 8
I53267
chorionic somatomammotropin-1 - rhesus macaque
C:Species: Macaca mulatta (rhesus macaque)
C>Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 16-Jul-1999
C:Accession: I53267
R:Golos, T.G.; Durning, M.; Fisher, J.M.; Fowler, P.D.
Endocrinology 133, 1744-1752, 1993
A:Title: Cloning of four growth hormone/chorionic somatomammotropin-related complementary DNAs from the placenta of rhesus macaques
A:Reference number: 153267; MID:94008724; PMID:8404617
A:Accession: I53267
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-217 <RES>
A:Cross-references: GB:L16552; NID:q293108; PID:AAA18839.1; PID:q293109
C:Superfamily: prolactin

Query Match 49.7%; Score 396; DB 2; Length 217;
Best Local Similarity 82.2%; Pred. No. 1.8e-31;
Matches 74; Conservative 11; Mismatches 5; Indels 0; Gaps 0;

QY 3 PPTPLSKLFDNAMI RAHRLHQLAFDTYQFEFEAYIPKQKYSFLQNPQTSLSSESISPT 62
DQ 26 PTPPLSKLFDNAMI RAHRLHQLAFDTYQFEFEAYIPKQKYSFLQNPQTSLSSESISPT 62
QY 63 SNREETQOKSNLELRISLLIQLSQWLEPVQ 92
DQ 86 SNREETQOKSNLELRISLLIQLSQWLEPVQ 117

RESULT 9
I67407
chorionic somatomammotropin A precursor [validated] - human
C:Species: Homo sapiens (man)
C>Date: 23-Oct-1981 #sequence_revision 23-Oct-1981 #text_change 08-Dec-2000
C:Accession: G32435; A94422; A93833; A93192; A90054; A94427; A61283; I55229;
R:Chen, E.Y.; Liang, Y.C.; Smith, D.H.; Barrera-Saldana, H.A.; Gelinas, R.E.; Seeburg
Endocrinology 47, 479-497, 1989
A:Title: The human growth hormone locus: nucleotide sequence, biology, and evolution
A:Reference number: A94422; MID:89307277; PMID:2744760
A:Accession: G32435
A:Molecule type: DNA
A:Residues: 1-217 <CHR>
A:Cross-references: GB:G30371; NID:q183148; PID:AAA52551.1; PID:q183151
R:Goodman, H.M.; Denoto, F.; Fiddes, J.C.; Halliwell, R.A.; Page, G.S.; Smith, S.;
In Mobilization and Reassembly of Genetic Information. Scott, W.A.; Werner, R.; Jose
A:Reference number: A94422
A:Accession: A94422
A:Molecule type: mRNA
A:Residues: 1-217 <GOO>
R:Ishihara, K.; Masuda, N.; Watahiki, M.; Yamakawa, M.; Shimizu, K.; Nagai, J.; Nakast
Biochem. Int. 16, 287-292, 1988
A:Title: cDNA cloning of human chorionic somatomammotropin-1 mRNA whose transcriptio
A:Reference number: I52342; MID:88209056; PMID:2835050
A:Accession: I52342
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-3 <IAN>
A:Cross-references: GB:M35419; NID:q506822
R:Shorwood, L.M.; Burstein, Y.; Schechter, I.
Proc. Natl. Acad. Sci. U.S.A. 76, 3819-3823, 1979
A:Title: Primary structure of the NH-2-terminal extra piece of the precursor to huma
A:Reference number: A93833; MID:80034970; PMID:291043
A:Accession: A93833
A:Molecule type: protein
A:Residues: 1-3-26 <SHE>
A:Experimental source: placenta
R:Shine, J.; Seeburg, P.H.; Martini, J.A.; Baxter, J.D.; Goodman, H.M.
Nature 270, 494-499, 1977
A:Title: Construction and analysis of recombinant DNA for human chorionic somatomamm
A:Reference number: A93192; MID:78071761; PMID:593368
A:Accession: A93192
A:Molecule type: DNA
A:Residues: 50-217 <SHI>
A:Experimental source: placenta
R:Li, C.H.; Dixon, J.S.; Chung, D.
Arch. Biochem. Biophys. 155, 95-110, 1973
A:Title: Amino acid sequence of human chorionic somatomammotropin.
A:Reference number: A90054; MID:73201971; PMID:4712450
A:Accession: A90054

```

[illegible]


```

A:Molecule type: protein
A:Residues: 1-190 <ZAK>
R:Zakin, M.M.; Poskus, E.; Dellacha, J.M.; Paladini, A.C.; Santome, J.A.
FEBS Lett. 34, 353-355, 1973
A:Title: The amino acid sequence of equine growth hormone.
A:Reference number: A91395; MID:74020362; PID:4747849
A:Accession: A91395
A:Molecule type: protein
A:Residues: 1-190 <ZAK>
R:Zakin, M.M.; Poskus, E.; Dellacha, J.M.; Paladini, A.C.; Santome, J.A.
FEBS Lett. 25, 77-82, 1972
A:Title: Amino acid sequences around the cystine residues in equine growth hormone.
A:Reference number: A91383
A:Accession: A91383
A:Molecule type: protein
A:Residues: 42-59:157-190 <ZA3>
R:Cliver, L.; Hartree, A.S.
Biochem. J. 109, 19-24, 1968
A:Title: Amino acid sequences around the cystine residues in horse growth hormone.
A:Reference number: A90240; MID:69368350; PID:4876100
A:Accession: A90240
A:Molecule type: protein
A:Residues: 175-190 <OLI>
C:Superfamily: prolactin
C:Keywords: hormone; pituitary
F:52-163,180-188/Disulfide bonds: #status experimental

Query Match 38.58; Score 306.5; DB 1; Length 190;
Best Local Similarity 65.21; Pred. No. 9.7e-23;
Matches 60; Conservative 14; Mismatches 17; Indels 1; Gaps 1;

QY 2 FPIPIPSRLFDNAMLRAHRLHQLAFDTYQCFEEAYIPKEQKYSFLQNPQTSLSFSSEIPT 61
DQ 1 FPAPLSSLSFANAVPAQHLHQLAANDYKKEERAYIPGORYS IQNAQAACFSETIPA 59

QY 62 PSNREETQQKSNLELLRISLLIOSWLEPVOL 93
DQ 60 PTKGDEAQQQSDMKLRAFLSLLIOSWLEGPVOL 91

RESULT 15
STMS
Somatotropin precursor - mouse
N:Alternate names: growth hormone
C:Species: Mus musculus (house mouse)
C>Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 28-May-1999
C:Accession: B23911
R:Blümler, C.J.H.; Tajamantes, F.
J. Biol. Chem. 260, 9574-9579, 1985
A:Title: Nucleotide sequence of mouse prolactin and growth hormone mRNAs and express
A:Reference number: A92548; MID:85261358; PID:2991252
A:Accession: B23911
A:Molecule type: mRNA
A:Residues: 1-216 <LIN>
A:CROSS-references: GR:X02891; GB:X03232; NID:g5i067; PIDN:CAA26650.1; PID:g5-065
C:Superfamily: prolactin
C:Keywords: anterior pituitary; growth factor; hormone
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-226/Product: somatotropin #status predicted <SIN>
F:78-163,206-214/Disulfide bonds: #status predicted

Query Match 38.28; Score 304.5; DB 1; Length 216;
Best Local Similarity 64.89; Pred. No. 1.8e-22;
Matches 59; Conservative 14; Mismatches 17; Indels 1; Gaps 1;

QY 2 FPIPIPSRLFDNAMLRAHRLHQLAFDTYQCFEEAYIPKEQKYSFLQNPQTSLSFSSEIPT 61
DQ 27 FPAPLSSLSFNAVLRQHLHQLAANDYKKEERAYIPGORYS IQNAQAACFSETIPA 85

QY 62 PSNREETQQKSNLELLRISLLIOSWLEPVQ 92
DQ 86 PTKGDEAQQQRTIMELLRFLSLLIOSWLEGPVQ 116

```

Search completed: September 16, 2003, 12:40:56
Job time : 18.5037 secs

GenCore version 5.0.6
Copyright (c) 1993 - 2003 GenPharm Inc.

OM protein - protein search, using sw mode

Run on: September 16, 2003, 12:33:32 : Search time 11.673.2 seconds
(without alignment)
504,234 Million cell updates/sec

Title: US-09-423-100-7

Perfect score: 797

Sequence: 1 MFPIPLSRLEFNAVLAHR.....IVKCCISGLYGVENON 59

Scoring table: BLSCM62

Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	451	57.8	217	1	SOMA_HUMAN
2	451	57.8	217	1	SOMA_PANTR
3	437	57.7	217	1	SOMA_MACRO
4	437	54.8	217	1	SOMA_SAIBR
5	434.5	54.5	217	1	SOMA_PANTR
6	432	54.2	217	1	SOMA_CALJA
7	426.5	53.5	217	1	SOMA_HUMAN
8	399	50.1	217	1	SOMA_MACRO
9	381	47.8	217	1	POL_HUMAN
10	310.5	39.0	216	1	SOMA_MESAU
11	307.5	38.6	190	1	SOMA_PALBO
12	306.5	38.5	216	1	SOMA_HORSE
13	306.5	38.5	217	1	SOMA_GALSE
14	306.5	38.5	217	1	SOMA_XYCPY
15	304.5	38.2	216	1	SOMA_XYFSE
16	302.5	38.0	216	1	SOMA_RABIT
17	302.5	37.8	216	1	SOMA_RAT
18	301.5	37.8	190	1	SOMA_COXAP
19	301.5	37.8	216	1	SOMA_CANPA
20	301.5	37.8	216	1	SOMA_FELCA
21	301.5	37.6	216	1	SOMA_FIG
22	299.5	37.6	216	1	SOMA_RUSVI
23	297.5	37.3	190	1	SOMA_LAMPA
24	295.5	37.1	190	1	SOMA_VULVU
25	291.5	36.6	215	1	SOMA_MONDO
26	291.5	36.6	215	1	SOMA_TRIVU
27	289.5	36.3	217	1	SOMA_BOVIN
28	289.5	36.3	217	1	SOMA_CEREL
29	289.5	36.3	217	1	SOMA_SHEEP
30	282.5	35.4	217	1	SOMA_RIBRJ
31	278.5	34.9	216	1	SOMA_MELGA
32	277.5	34.8	110	1	INS_CRAE
33	277.5	34.8	110	1	INS_RABIT

34	275.5	34.6	216	1	SOMA_CHICK
35	274.5	34.4	217	1	SOMA_STRCA
36	273.5	34.3	51	1	INS_BALPH
37	273.5	34.3	51	1	INS_ELENA
38	273.5	34.3	110	1	INS_MACFA
39	272.5	34.2	190	1	SOMA_CRONO
40	272.5	34.2	190	1	INS_PANTR
41	270	33.9	110	1	INS_HUMAN
42	268.5	33.7	51	1	INS_ACOCA
43	258.5	33.7	291	1	SOMA_CHEMY
44	256.5	33.4	105	1	INS_BOVIN
45	246.5	33.4	110	1	INS_SPETR

ALIGNMENTS

Result 1

TO	SOMA_HUMAN	STANDARD	PRT:	217 AA
AC	P01241: Q14405; 016631; Q9HB21; Q9UMJ7; Q9UNL5;			
BT	22-JUN-1986 (Rel. 0); Created			
CT	01-MAR-1992 (Rel. 2); Last sequence update;			
DT	15-SEP-2003 (Rel. 42; Last annotation update)			
DE	Somatotropin precursor (Growth hormone) (GH) (GH-N) (Pituitary growth hormone) (Growth hormone 1).			
DS	GH1.			
GN	Homo sapiens (Human).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	(1)			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RA	MEDLINE-80034477; PubMed-386281;			
RT	Roskam W., Rougeon F.;			
FT	"Molecular cloning and nucleotide sequence of the human growth hormone structural gene."			
RL	Nucleic Acids Res. 7:305-320(1979).			
RN	(2)			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RA	MEDLINE-79203293; PubMed-377456;			
RT	Marital J.A., Halliwell R.A., Baxter J.D., Goodman H.M.;			
FT	"Human growth hormone: complementary DNA cloning and expression in bacteria."			
RL	Science 205:602-607(1979).			
RN	(3)			
RP	SEQUENCE FROM N.A. (ISOFORM 1), AND POSSIBLE ALTERNATIVE SPLICING.			
RA	MEDLINE-82014939; PubMed-6256091;			
RT	Denoto F.N., Moore D.B., Goodman H.M.;			
FT	"Human growth hormone DNA sequence and mRNA structure: possible alternative splicing."			
RL	Nucleic Acids Res. 9:3719-3730(1981).			
RN	(4)			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE-83182010; PubMed-7169009;			
RT	Seeburg P.H.;			
FT	"The human growth hormone gene family: nucleotide sequences show recent divergence and predict a new polypeptide hormone."			
RL	DNA 1:239-245(1982).			
RN	(5)			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE-89307277; PubMed-2744760;			
RT	Chen E.Y., Liao Y.G., Smith D.H., Barrera-Saldana H.A.;			
FT	Gellinas R.E., Seeburg P.H.;			
RL	"The human growth hormone locus: nucleotide sequence, biology, and evolution."			
RN	Genomics 4:479-497(1989).			
RP	SEQUENCE FROM N.A. (ISOFORM 3).			
RA	TISSUE-Pituitary;			
RT	Gu J., Huang Q.-H., Li N., Xu S.-H., Han Z.-G., Fu G., Chen Z.;			
FT	"A novel gene expressed in human pituitary."			
RL	Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.			

RA Lewis J.J., Borewald L.F., Lewis L.J.;
 RT "The 20,000-dalton variant of human growth hormone: location of the
 PT amino acid deletions";
 RL Biochem. Biophys. Res. Commun. 92:511-516(1980).
 RA [17]
 RP DEAMINATION OF GDN-163 AND ASN-178.
 RX MEDLINE=82052997; PubMed=7028740;
 RA Lewis J.J., Singh R.N., Bonewald L.F., Seavey B.K.;
 RT "Altered proteolytic cleavage of human growth hormone as a result of
 PT deamidation";
 RL J. Biol. Chem. 256:11645-11650(1981).
 RA [18]
 RP REVIEW.
 RX MEDLINE=99321812; PubMed=10393484;
 RA Baumann G.;
 RT "Growth hormone heterogeneity in human pituitary and plasma";
 RL Horm. Res. 51 Suppl. 1:2-6(1999).
 RA [19]
 RP 3D-STRUCTURE MODELING.
 RX MEDLINE=88190073; PubMed=3447173;
 RA Cohen F.E., Kuntz I.D.;
 RT "Prediction of the three-dimensional structure of human growth
 PT hormone";
 RL Proteins 2:162-166(1987).
 RA [20]
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
 RX MEDLINE=92196577; PubMed=1549776;
 RA de Vos A.M., Uitsch M., Kossiakoff A.A.;
 RT "Human growth hormone and extracellular domain of its receptor:
 PT crystal structure of the complex";
 RL Science 255:306-312(1992).
 RA [21]
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
 RX MEDLINE=95075452; PubMed=794244;
 RA Somers W., Uitsch M., de Vos A.M., Kossiakoff A.A.;
 RT "The X-ray structure of a growth hormone-prolactin receptor complex";
 RL Nature 372:476-481(1994).
 RA [22]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RX MEDLINE=97113023; PubMed=8943276;
 RA Pavlovsk A.G., Wisdawa A.;
 RT "The crystal-structure of wild-type growth-hormone at 2.5-A
 PT resolution";
 RL Protein Pept. Lett. 2:333-340(1995).
 RA [23]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RX MEDLINE=97113023; PubMed=8943276;
 RA Sundstrom M., Lundqvist T., Roedin J., Giebel L.B., Milligan D.,
 RT "Crystal structure of an antagonist mutant of human growth hormone,
 PT G20K, in complex with its receptor at 2.9-A resolution";
 RL J. Mol. Chem. 27:32197-32203(1996).
 RA [24]
 RP VARIANT CYS-105.
 RX MEDLINE=99318093; PubMed=10391209;
 RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.O.,
 RA Lander E.S.;
 RT "Characterization of single-nucleotide polymorphisms in coding regions
 PT of human genes";
 RL Nat. Genet. 22:231-238(1999).
 RA [25]
 RP ERRATUM.
 RX MEDLINE=99318093; PubMed=10391209;
 RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.O.,
 RA Lander E.S.;
 RT "Characterization of single-nucleotide polymorphisms in coding regions
 PT of human genes";
 RL Nat. Genet. 22:231-238(1999).
 RA [26]
 RP FUNCTION: Plays an important role in growth control. Its major
 RT role in stimulating body growth is to stimulate the liver and
 CC other tissues to secrete IGF-1. It stimulates both the
 CC differentiation and proliferation of myoblasts. It also stimulates
 CC

RA [7]
 RP SEQUENCE FROM N.A. (ISOFORM 4).
 RC TISSUE=pituitary;
 RX MEDLINE=20402571; PubMed=10941946;
 RA Hu Y.-J., Huang C.-H., Li Y.-H., Jiang C.-L., Fu S., Zhang Q.-H.,
 RA Gu B.-W., Dai M., Mo Y.-F., Gao G.-F., Rong R., Ye M., Zhou J.,
 RA Xu S.-H., Gu J., Shi J.-X., Jin W.-R., Zhang C.-K., Wu T.-M.,
 RA Huang G.-Y., Chen Z., Chen M.-D., Chen J.-L.;
 RT "Gene expression profiling in the human hypothalamus-pituitary-adrenal
 PT axis and full-length cDNA cloning";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:9543-9548(2000).
 RA [8]
 RP SEQUENCE OF 1-26 FROM N.A.
 RX MEDLINE=8637393; PubMed=3912261;
 RA Gray G.L., Baldridge J.S., McKee K.S., Heynecker H.L., Chang C.N.;
 RT "Periplasmic production of correctly processed human growth hormone in
 PT Escherichia coli: natural and bacterial signal sequences are
 RT interchangeable";
 RL Gene 99:247-254(1985).
 RA [9]
 RP SEQUENCE OF 27-217.
 RX MEDLINE=69289202; PubMed=6810834;
 RA Li C.H., Dixon J.S., Liu W.K.;
 RT "Human pituitary growth hormone. XIX. The primary structure of the
 PT hormone";
 RL Arch. Biochem. Biophys. 131:70-91(1969).
 RA [10]
 RP SEQUENCE OF 27-217, AND REVISIONS.
 RX MEDLINE=72143935; PubMed=5144027;
 RA Li C.H., Dixon J.S.;
 RT "Human pituitary growth hormone. 32. The primary structure of the
 PT hormone: revision";
 RL Arch. Biochem. Biophys. 146:233-236(1971).
 RA [11]
 RP REVISION.
 RX MEDLINE=73092028; PubMed=4675454;
 RA Bewley T.A., Dixon J.S., Li C.H.;
 RT "Sequence comparison of human pituitary growth hormone, human
 PT chorionic somatomammotropin, and ovine pituitary growth and
 RT lactogenic hormones";
 RL Int. J. Pept. Protein Res. 4:261-267(1972).
 RA [12]
 RP SEQUENCE OF 27-61 AND 102-124.
 RX MEDLINE=71139765; PubMed=5229045;
 RA Niall H.D.;
 RT "Revised primary structure for human growth hormone";
 RL Nature New Biol. 230:90-91(1971).
 RA [13]
 RP REVISIONS TO 119-120 AND 157-159.
 RX MEDLINE=71151958; PubMed=5229528;
 RA Niall H.D., Hogan M.L., Sauter R., Rosenblum I.V., Greenwood F.C.;
 RT "Sequences of pituitary and placental lactogenic and growth hormones:
 PT evolution from a primordial peptide by gene reduplication";
 RL Proc. Natl. Acad. Sci. U.S.A. 68:866-869(1971).
 RA [14]
 RP REVISION.
 RX MEDLINE=81171361; PubMed=7462247;
 RA Chapman G.B., Rogers K.M., Erttaint T., Bradshaw R.A., Bates G.J.,
 RA Turner C., Garry P.D., Crane-Robinson C.;
 RT "The 20,000 molecular weight variant of human growth hormone.
 PT Preparation and some physical and chemical properties";
 RL J. Biol. Chem. 256:3395-2401(1981).
 RA [16]
 RP SEQUENCE OF 46-80 (ISOFORM 2).
 RX MEDLINE=80130196; PubMed=7356479;

```

Query Match      57.8% Score 461; DB 1; Length 217;
Best Local Similarity 70.3%; Pred. No. 4; Gap 36;
Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 2 FFTIPSLRFDNMLRAHRLHOLAFDTYQEFEEAYIPKEQKYSFLQNFQNTSCFSESIP1 86
DE 2 FFTIPSLRFDNMLRAHRLHOLAFDTYQEFEEAYIPKEQKYSFLQNFQNTSCFSESIP1 86
QY 27 FFTIPSLRFDNMLRAHRLHOLAFDTYQEFEEAYIPKEQKYSFLQNFQNTSCFSESIP1 86
DE 27 FFTIPSLRFDNMLRAHRLHOLAFDTYQEFEEAYIPKEQKYSFLQNFQNTSCFSESIP1 86
QY 62 PSNRRETOOKSNLELRISLLLIQSWLEPVCVGTGRTFVNORHGS-----HIVE 111
DE 62 PSNRRETOOKSNLELRISLLLIQSWLEPVCVGTGRTFVNORHGS-----HIVE 111
QY 87 PSNRRETOOKSNLELRISLLLIQSWLEPVCVGTGRTFVNORHGS-----HIVE 111
DE 87 PSNRRETOOKSNLELRISLLLIQSWLEPVCVGTGRTFVNORHGS-----HIVE 111
QY 112 ALYLWCG--ERGFYTPKTRGIVEQ 134
DE 112 ALYLWCG--ERGFYTPKTRGIVEQ 134
QY 146 GTQLNGRLEDC---SPRTGQIFKQ 167
DE 146 GTQLNGRLEDC---SPRTGQIFKQ 167

RESULT 2
SOMA_PANTR
ID SOMA_PANTR STANDARD: PRT: 217 AA.
AC PS9756;
DT 28-FEB-2003 (Rel. 4); Created)
DT 28-FEB-2003 (Rel. 4); Last sequence update)
DT 28-FEB-2003 (Rel. 4); Last annotation update)
DE Somatotropin precursor (Growth hormone) (GH) (GH-N) (Pituitary growth hormone) (Growth hormone 1).
GN GH1.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP Revol A., Esquivel D., Santiago D., Barrera-Saidana H.;
RT "Independent duplication of the growth hormone gene in three Anthropoid lineages."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Plays an important role in growth control. Its major role in stimulating body growth is to stimulate the liver and other tissues to secrete IGF-1. It stimulates both the differentiation and proliferation of myoblasts. It also stimulates amino acid uptake and protein synthesis in muscle and other tissues (by similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announcement/ or send an email to license@sib-sib.ch.)
CC EMBL: AF374232; AML72284.1;
CC InterPro: IPR001400; Somatotropin.
CC Pfam: PF00103; hormone; 1.
CC PRINIS: PK00836; SOMATOTROPIN.
CC PROSITE: PS00266; SOMATOTROPIN_1; 1.
CC PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW Hormone; Pituitary; Signal.
FT SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 27 217 SOMATOTROPIN.
FT DISULFID 79 191 BY SIMILARITY.
FT DISULFID 208 215 BY SIMILARITY.
FT CONFLICT 100 100 E -> Q (IN REF. 2).
FT CONFLICT 179 175 N -> D (IN REF. 2).
SQ SEQUENCE 217 AA; 24843 MW; FEA295DEDE9E18674 CRG64;

Query Match      57.8% Score 461; DB 1; Length 217;
Best Local Similarity 70.3%; Pred. No. 4; Gap 36;
Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 2 FFTIPSLRFDNMLRAHRLHOLAFDTYQEFEEAYIPKEQKYSFLQNFQNTSCFSESIP1 86
DE 2 FFTIPSLRFDNMLRAHRLHOLAFDTYQEFEEAYIPKEQKYSFLQNFQNTSCFSESIP1 86
QY 27 FFTIPSLRFDNMLRAHRLHOLAFDTYQEFEEAYIPKEQKYSFLQNFQNTSCFSESIP1 86
DE 27 FFTIPSLRFDNMLRAHRLHOLAFDTYQEFEEAYIPKEQKYSFLQNFQNTSCFSESIP1 86
QY 62 PSNRRETOOKSNLELRISLLLIQSWLEPVCVGTGRTFVNORHGS-----HIVE 111
DE 62 PSNRRETOOKSNLELRISLLLIQSWLEPVCVGTGRTFVNORHGS-----HIVE 111
QY 87 PSNRRETOOKSNLELRISLLLIQSWLEPVCVGTGRTFVNORHGS-----HIVE 111
DE 87 PSNRRETOOKSNLELRISLLLIQSWLEPVCVGTGRTFVNORHGS-----HIVE 111
QY 112 ALYLWCG--ERGFYTPKTRGIVEQ 134
DE 112 ALYLWCG--ERGFYTPKTRGIVEQ 134
QY 146 GTQLNGRLEDC---SPRTGQIFKQ 167
DE 146 GTQLNGRLEDC---SPRTGQIFKQ 167

RESULT 3
SOMA_MACHU
ID SOMA_MACHU STANDARD: PRT: 217 AA.
AC P13023;
DT 01-OCT-1993 (Rel. 27; Created)
DT 01-OCT-1994 (Rel. 30; Last sequence update)
DT 28-FEB-2003 (Rel. 4); Last annotation update)
DE Somatotropin precursor (Growth hormone) (GH) (GH-N) (Pituitary growth hormone) (Growth hormone 1).
GN GH1.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RP MEDLINE=94008724; PubMed=8404617;
RA Golos I.G., Dunning M., Fisher J.M., Fowler P.D.;
RT "Cloning of four growth hormone/chorionic somatomammotropin-related complementary deoxyribonucleic acids differentially expressed during pregnancy in the rhesus monkey placenta."
RL Endocrinology 133:1744-1752(1993).
RN [2]
RP SEQUENCE OF 27-217.
RP MEDLINE=86129460; PubMed=3080959;
RA Li C.H., Chung D., Lalm H.W., Stein S.;
RT "The primary structure of monkey pituitary growth hormone."
RA Arch. Biochem. Biophys. 245:287-291(1986).
CC -!- FUNCTION: Plays an important role in growth control. Its major role in stimulating body growth is to stimulate the liver and other tissues to secrete IGF-1. It stimulates both the differentiation and proliferation of myoblasts. It also stimulates amino acid uptake and protein synthesis in muscle and other tissues.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announcement/ or send an email to license@sib-sib.ch.)
CC EMBL: L16556; AAA18842.1;
CC Pfam: PF00103; hormone; 1.
CC RASP: P02241; IAXI.
CC InterPro: IPR001400; Somatotropin.
CC Pfam: PF00103; hormone; 1.
CC PROSITE: PS00266; SOMATOTROPIN_1; 1.
CC PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW Hormone; Pituitary; Signal.
FT SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 27 217 SOMATOTROPIN.
FT DISULFID 79 191 BY SIMILARITY.
FT DISULFID 208 215 BY SIMILARITY.
FT CONFLICT 100 100 E -> Q (IN REF. 2).
FT CONFLICT 179 175 N -> D (IN REF. 2).
SQ SEQUENCE 217 AA; 24913 MW; 2C5180341EEC46D0 CRG64;

```

[illegible]

```

DT 28-FEB-2003 (Rel. 41, Last annotation update)
DR Somatotropin precursor (Growth hormone)
GN GHI
OS Callithrix jacchus (Common marmoset)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
OC Callithrix
CX NCBI_TaxID=9483;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RA Wallis O.C., Wallis M.J.
RT "Cloning and characterization of a relative growth hormone encoding
RL gene from the marmoset (Callithrix jacchus)".
RL Submitted (AUG-2003) to the EMBL/GenBank/CCDB databases.
CC 1- FUNCTION: Plays an important role in growth control. Its major
CC role in stimulating body growth is to stimulate the liver and
CC other tissues to secrete IGF-1. It stimulates both the
CC differentiation and proliferation of myoblasts. It also stimulates
CC amino acid uptake and protein synthesis in muscle and other
CC tissues (by similarity).
CC 1- SUBCELLULAR LOCATION: Secreted.
CC 1- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL, in collaboration
CC with the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-ch).
DR EMBL: AC297563; CAC03481.1;
DR HSSP: P01241; 1A22.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone_1.
DR PROSITE: PS00266; SOMATOTROPIN_1: 1.
DR PROSITE: PS00338; SOMATOTROPIN_2: 1.
KW Hormone; Pituitary; Signal.
FT SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 27 217 SOMATOTROPIN.
FT DISULFID 79 191 BY SIMILARITY.
FT DISULFID 208 215 BY SIMILARITY.
SQ SEQUENCE 217 AA: 24955 MR: E10251A1CEA192 CRG04:
Query Match 54.2%; Score 132; DS 1; Length 217;
Best local Similarity 91.2%; Pred. No. 3,26-16;
Matches #3; Conservative 5; Mismatches 4; Gaps 0;
QY 2 FTPTLSRLEFDNMLRAHPLRLKADIVQFFETATPKQKYSPLQNTQKTSFESPT 41
Db 27 FTPTLSRLEFDNMLRAHPLRLKADIVQFFETATPKQKYSPLQNTQKTSFESPT 40
QY 62 PSNREETQCKSNLELRISLLELQSWLRFVQ 52
Db 87 PASKKETQCKSNLELRISLLELQSWLRFVQ 117
RESULT 7
SOM2_HUMAN STANDARD; iRT: 217 AA.
AC P01242; PC0587;
DT 21-JUL-1986 (Rel. 01, Created)
DI 28-FEB-2003 (Rel. 41, Last sequence update)
DE Growth hormone variant precursor (GH-V) (Placenta-specific growth
DE hormone) (Growth hormone 2).
GN GH2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea;
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).

```

```

FX MEDLINE-83182010; PubMed-7169009;
FA Seeburg P.H.;
AT "The human growth hormone gene family: nucleotide sequences show
A recent divergence and predict a new polypeptide hormone.";
AN DNA 1:239-245(1982).
AX [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
FX MEDLINE-86243769; PubMed-4379597;
RA Cooke N.E., Hay J., Emery C.G., Liebhauer S.A.;
RT "Two distinct species of human growth hormone-variant mRNA in the
PT human placenta predict the expression of novel growth hormone
PT proteins";
PL J Biol. Chem. 265:9001-9006(1988).
RX [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
FX MEDLINE-89021584; PubMed-2460050;
RA Iqbal A., Scippo M.L., Frankenne F., Hennen G.;
RT "Cloning and nucleotide sequence of placental hGH-V cDNA.";
AN Arch. Int. Physiol. Biochim. 95:63-67(1988).
AX [4]
RP SEQUENCE FROM N.A.
FX MEDLINE-89307277; PubMed-2744760;
RA Chen E.Y., Liao Y.C., Smith D.H., Barrera-Saldana H.A.;
RT "The human growth hormone locus: nucleotide sequence, biology, and
PT evolution.";
PL Genomics 4:479-497(1989).
RX [5]
RP SEQUENCE FROM N.A.
FX TISSUE=Placenta;
FX MEDLINE-22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan B., Moore L., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rata S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gutarane P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Viatrala D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Maman A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M., Skalska U., Smailus D.E.,
RA Schurch A., Schein J.B., Jones S.J.N., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
PT human and mouse cDNA sequences.";
AN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
AX [6]
RP REVIEW
FX MEDLINE-99321812; PubMed-10393484;
FA Baumann G.;
RT "Growth hormone heterogeneity in human pituitary and plasma.";
AN Horm. Res. 51 Suppl. 1:2-6(1999).
AX 1- FUNCTION: Plays an important role in growth control. Its major
AX role in stimulating body growth is to stimulate the liver and
AX other tissues to secrete IGF-1. It stimulates both the
AX differentiation and proliferation of myoblasts. It also stimulates
AX amino acid uptake and protein synthesis in muscle and other
AX tissues.
AX 1- SUBUNIT: Monomer, dimer, trimer, tetramer and pentamer, disulfide-
AX linked or non-covalently associated, in homopolymeric and
AX heteropolymeric combinations. Can also form a complex either with
AX GHBP or with the alpha2-macroglobulin complex.
AX 1- SUBCELLULAR LOCATION: Secreted.
AX 1- ALTERNATIVE PRODUCTS:
AX Name=1; Synonyms=GH-V1;
AX IsoId=P01242-1; Sequence=Displayed;
AX Name=2; Synonyms=GH-V2;

```


DT 01-APR-1988 (Rel. 07, last sequence update)
 DI 15-SEP-2003 (Rel. 42, last annotation update)
 DE Lactogen precursor (Chorionamniototropin) (Chorionic somatomammotropin)
 GN CSH1 AND CSH2.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP MEDLINE-85033426; PubMed-6208192;
 RA Selby M.J., Barta A., Baxter J.D., Bell G.L., Eberhardt K.L.,
 RT "Analysis of a major human chorionic somatomammotropin gene: Evidence
 RT for two functional promoter elements."
 RL J. Biol. Chem. 259:13131-13138(1984).
 RN [2]
 RP SEQUENCE FROM N.A. (GENE CSH1).
 RX MEDLINE-87161235; PubMed-3010560;
 RA Hirt H., Kimmel J., Birnbaum M.J., Chen R.Y., Seeburg P.H.,
 RA Eberhardt K.L., Barta A.,
 RT "The human growth hormone gene locus: structure, evolution, and
 RT allelic variations."
 RL DNA 6:59-70(1987).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-83160916; PubMed-6300255;
 RA Barrera-Saldana H.A., Seeburg P.H., Saunders G.F.,
 RT "Two structurally different genes produce the same secreted human
 RT placental lactogen hormone."
 RL J. Biol. Chem. 258:3787-3793(1983).
 RN [4]
 RP SEQUENCE FROM N.A. (GENES CSH1 AND CSH2).
 RX MEDLINE-89507277; PubMed-2744760;
 RA Choe E.Y., Liao Y.C., Smith D.H., Barrera-Saldana H.A., Gellius R.E.,
 RA Seeburg P.H.,
 RT "The human growth hormone locus: nucleotide sequence, biology, and
 RT evolution."
 RL Genomics 4:479-497(1989).
 RN [5]
 RP SEQUENCE.
 RX MEDLINE-83182010; PubMed-7159009;
 RA Seeburg P.H.,
 RT "The human growth hormone gene family: nucleotide sequences show
 RT recent divergence and predict a new polypeptide hormone."
 RL DNA 1:239-245(1982).
 RN [6]
 RP SEQUENCE FROM N.A.
 CC TISSUE-Placenta, and Uterus.
 RX MEDLINE-22382357; PubMed-12477932;
 RA Strausberg K.L., Feingold E.A., Grouse L.H., Borge J.G.,
 RA Klausner R.D., Collins F.S., Wagner T., Shenmen S.M., Schaller G.D.,
 RA Altschul S.F., Zeeberg B., Bietow K.H., Schaefer G.F., Blatt N.K.,
 RA Hopkins R.P., Jordan H., Moore T., Wax S., Wang J., Hsieh P.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Abert M.,
 RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Schaefer J.P.,
 RA Brownstein M.J., Udgin T.R., Toshcyki S., Cantarel D., Prange C.,
 RA Raha S., Loquellane N.A., Peters G.J., Andrews E.P., Hutchy S.H.,
 RA Bosak S.A., McKernan K.J., Malek S.A., Samarasekera D.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Guy J.J., Hilyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Rahey J., Heitron E., Kettner M., Madan A., Kodumuru R., Sanchez A.,
 RA Whiting M., Madan A., Yount A.C., Shrivastava Y., Bontirud G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield V.S.N., Krzywicki M.I., Skalska U., Smalios D.E.,
 RA Schnorcher A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [7]
 RP SEQUENCE OF 50-217 FROM N.A.
 RX MEDLINE-78071761; PubMed-593368;
 RA Shine J., Seeburg P.H., Martel J.A., Baxter J.D., Goodman H.M.,

PT "Construction and analysis of recombinant DNA for human chorionic
 PT somatomammotropin."
 RC Nature 270:494-499(1977).
 RN [8]
 RP SEQUENCE OF 27-217.
 RX MEDLINE-77320197; PubMed-4712450;
 RA Li C.H., Dixon J.S., Cheng D.,
 RT "Amino acid sequence of human chorionic somatomammotropin."
 RL Arch. Biochem. Biophys. 155:95-110(1973).
 RN [9]
 RP SEQUENCE OF 27-117.
 RX MEDLINE-72016133; PubMed-5286363;
 RA Sherwood L.M., Handwerger S., McLaurin W.D., Lannier M.,
 RT "Amino acid sequence of human placental lactogen."
 RL Nature New Biol. 233:59-61(1971).
 RN [10]
 RP ERRATUM.
 RA Sherwood L.M., Handwerger S., McLaurin W.D., Lannier M.,
 RL Nature New Biol. 235:54-54(1972).
 RN [11]
 RP INTERCHAIN DISULFIDE BONDS.
 RX MEDLINE-79173081; PubMed-438159;
 RA Schneider A.R., Kowalski K., Russell J., Sherwood L.M.,
 RT "Identification of the interchain disulfide bonds of dimeric human
 RT placental lactogen."
 RL J. Biol. Chem. 254:3782-3787(1979).
 CC [1] FUNCTION: SIMILAR TO THAT OF SOMATOTROPIN.
 CC [2] SUBCELLULAR LOCATION: Secreted.
 CC [3] MISCELLANEOUS: THE SEQUENCE OF CSH1 IS SHOWN.
 CC [4] SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMFL: V00573; CAA23836.1;
 DR EMFL: J00289; AAA98747.1;
 DR EMFL: K02401; AAA52115.1;
 DR EMFL: M55894; AAA52116.1;
 DR EMFL: J03071; AAA52551.1;
 DR EMFL: J00118; AAA38621.1;
 DR EMFL: R002717; AAH02717.1;
 DR EMFL: BC005921; AAH05921.1;
 DR EMFL: BC028756; AAH20756.1;
 DR PIR: A26449; A26449.
 DR PIR: C32435; ICHUG.
 DR HSP: P01241; 1A22.
 DR Genew: HGNC:2440; CSH1.
 DR Genew: HGNC:2441; CSH2.
 DR MIM: 150200;
 DR GO: GO:0007565; P:pregnancy; TAS.
 DR InterPro: IPR001400; Somatomotropin.
 DR Pfam: PF00103; hormone; 1.
 DR PROSITE: PS0266; SOMATOTROPIN_1; 1.
 DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
 KW Hormone; Placenta; Multigene family; Signal.
 FT SIGNAL: 1 26
 FT CHAIN: 27 217 LACTOGEN.
 FT DISULFID: 79 191
 FT DISULFID: 208 215
 FT DISULFID: 208 208 INTERCHAIN (WITH C-215 IN A DIMER).
 FT DISULFID: 215 215 INTERCHAIN (WITH C-208 IN A DIMER).
 FT VARIANT: 3 3 P -> A (IN CSH2).
 FT VARIANT: 104 105 /FTID-VAR_007166.
 FT VARIANT: 84 84 IS -> L (IN CSH2).
 FT CONFLICT: 84 84 I -> T (IN REF. 9).
 FT CONFLICT: 95 95 MISSING (IN REF. 9).
 FT CONFLICT: 116 116 MISSING (IN REF. 9).

```

FT CONFLICT 134 135 SUD -> BBS (IN REF. 9).
SQ SEQUENCE 217 AA: 25020 MW: 235BDDCA7A713F43 CRC64:

Query Match: 47.8% Score 381; DB 1; Length 217;
Best Local Similarity 82.0% Pred. No. 3-26-30;
Matches 73; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 4 TPLSRFLDNNMLRAHKLHJADTYOFFBEAVIPKHKYKSYFLGNKGTSLNPSFSTPLS 63
DB 29 TVPLSRFLDNNMLRAHKLHJADTYOFFBEAVIPKHKYKSYFLGNKGTSLNPSFSTPLS 68
QY 64 NRETOOKSNLELLRSLLLQSLWLPVQ 92
DB 89 NRETOOKSNLELLRSLLLQSLWLPVQ 117

RESULT 10
SOMA_MESAU STANDARD; PRT: 216 AA.
AC P3786;
DT 01-OCT-1994 (Rel. 30, Create)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 4, Last annotation update)
DE Somatotropin precursor (Growth hormone).
GN GH1 OR GH.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Grantata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetidae;
OC Mesocricetus.
CN NCBL_TaxID=10046;
RX NCBI_1
RP SEQUENCE FROM N.A.
RX MEDLINE-92063650; PubMed-195488;
RA Southard J.N., Sanchez-Jimenez P., Campbell G.T., Zalamantes P.
RT "Sequence and expression of hamster prolactin and growth hormone messenger RNAs".
RL Endocrinology 129:2365-2371(1991).
CC -!- FUNCTION: Plays an important role in growth control. Its major role in stimulating body growth is to stimulate the liver and other tissues to secrete IGF-1. It stimulates both the differentiation and proliferation of myoblasts. It also stimulates amino acid uptake and protein synthesis in muscle and other tissues.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/agreement/ or send an email to license@sb-sib.ch).
CC
CC EMBL: S66299; AAB20368.1;
CC PIR: B49159; B49159.
CC HSP: P01246; IBSI.
CC InterPro: IPR001400; Somatotropin.
CC Pfam: PF00103; Hormone_1.
CC PROSITE: PS00265; SOMATOTROP_N_1; 1.
CC PROSITE: PS00138; SOMATOTROP_N_2; 1.
CC Hormone; Pituitary; Signal.
KW SIGNAL
FT SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 27 216 SOMATOTROPIN.
FT DISULFID 78 189 BY SIMILARITY.
FT DISULFID 206 214 BY SIMILARITY.
SQ SEQUENCE 216 AA: 24590 MW: 2569CE32AB6F1168 CRC64:

Query Match 39.0% Score 310.5; DB 1; Length 216;
Best Local Similarity 67.0% Pred. No. 2.5e-23;
Matches 61; Conservative 11; Mismatches 16; Indels 1; Gaps 1;

Oy 2 FPLSLSELNAMLRAHRLHQLAFDYQFEAVIPKFKYKSYFLGNKGTSLNPSFSTPLS 61

```

DE Somatotropin precursor (Growth hormone).
GN GH:
OS Equus caballus (Horse).
OC Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary;
RX MEDLINE=94266171; PubMed=8236352;
RA Ascacio-Martinez J.A., Barrera-Salazar H.A.
RT "Sequence of a cDNA encoding horse growth hormone."
RL Gene 143:299-300(1994).
RN [2]
RP SEQUENCE OF 27-216.
RX MEDLINE=77005410; PubMed=965151;
RA Zakin M.M., Poskus E., Lantton A.A., Ferrara J., Santome J.A.,
RA Bellacha J.M., Paladini A.C.
RT "Primary structure of equine growth hormone."
RL Int. J. Pept. Protein Res. 3:435-444(1976).
RN [3]
RP PRELIMINARY SEQUENCE OF 27-216.
RX MEDLINE=74020362; PubMed=4747845;
RA Zakin M.M., Poskus E., Bellacha J.M., Paladini A.C., Santome J.A.,
RT "The amino acid sequence of equine growth hormone."
RL FEBS Lett. 34:353-355(1973).
RN [4]
RP SEQUENCE OF 69-95 AND 183-216.
RA Zakin M.M., Poskus E., Bellacha J.M., Paladini A.C., Santome J.A.,
RI "Amino acid sequences around the cystine residues in equine growth hormone."
RL FEBS Lett. 25:77-82(1972).
RN [5]
RP SEQUENCE OF 202-216.
RX MEDLINE=68368390; PubMed=4676100;
RA Oliver L., Hartree A.S.
RI "Amino acid sequences around the cystine residues in horse growth hormone."
RL Biochem. J. 109:19-24(1968)
CC -!- FUNCTION: plays an important role in growth control. Its major role in stimulating body growth is to stimulate the liver and other tissues to secrete IGF-1. It stimulates both the differentiation and proliferation of myoblasts. It also stimulates amino acid uptake and protein synthesis in muscle and other tissues.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@sib-sib.ch).
CC EM36; 002929; AAA21027.1;
DR HSP; F01246; IBS1.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone; 1.
DR PRINTS; PR00836; SOMATOTROPIN.
DR PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW Hormone; Pituitary; Signal.
CC SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 27 216 SOMATOTROPIN.
FT DISULFID 78 159
FT DISULFID 206 214
CC SEQUENCE 216 AA; 37ABS73394C1AC CRC64;
Query Match: 38.5%; Score 306.5; DB 1; Length 216;
Best local similarity 65.2%; Pred. No. 6, 20-23;
Matches 60; Conservative 14; Mismatches 17; Indels 1; Gaps 1;

QY 2 FPTPLSLRFDNAKLRAHRLHQLAFDTYQEFERAYIPKEOKYSLQNPQTSLSFSISPT 61
DQ 27 FPAMPUSLFFANVLAQHLHQLAADTYKEFERAYIPGGRYS-IQNAAFCFSETIPA 85
QY 62 PSNREETOOKSNLELRISLLIQSWLGPVOL 93
DQ 86 FTGKDEAQRSDMELLRFSLLIQSWLGPVOL 117
RESID 13
IC SOMA_GA:SE STANDARD; PRI; 217 AA.
AC CCKAL;
RT 28-FEB-2003 (Rel. 4); Created;
RT 28-FEB-2003 (Rel. 4); Last sequence update;
RT 28-FEB-2003 (Rel. 4); Last annotation update;
DE Somatotropin precursor (Growth hormone).
GN GH:
OS Galago senegalensis (Northern lesser bushbaby).
OC Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Strepsirhini; Galagonidae; Galago.
OX NCBI_TaxID=9465;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=11141192;
RA Adkins R.M., Nekrutenko A., Li W.-H.
RT "Bushbaby growth hormone is much more similar to nonprimate growth hormones than to rhesus monkey and human growth hormones."
RL Mol. Biol. Evol. 18:55-61(2001).
CC -!- FUNCTION: Plays an important role in growth control. Its major role in stimulating body growth is to stimulate the liver and other tissues to secrete IGF-1. It stimulates both the differentiation and proliferation of myoblasts. It also stimulates amino acid uptake and protein synthesis in muscle and other tissues.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@sib-sib.ch).
CC EM36; AF292938; AAC44952.1;
DR HSP; F01246; IBS1.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone; 1.
DR PRINTS; PR00836; SOMATOTROPIN.
DR PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW Hormone; Pituitary; Signal.
CC SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 27 217 SOMATOTROPIN.
FT DISULFID 79 190
FT DISULFID 207 215
CC SEQUENCE 217 AA; 2F616C031136F005 CRC64;
Query Match: 38.5%; Score 306.5; DB 1; Length 217;
Best local similarity 55.2%; Pred. No. 6, 20-23;
Matches 60; Conservative 14; Mismatches 17; Indels 1; Gaps 1;

QY 2 FPTPLSLRFDNAKLRAHRLHQLAFDTYQEFERAYIPKEOKYSLQNPQTSLSFSISPT 61
DQ 27 FPAMPUSLFFANVLAQHLHQLAADTYKEFERAYIPGGRYS-IQNAAFCFSETIPA 86
QY 62 PSNREETOOKSNLELRISLLIQSWLGPVOL 93
DQ 87 FTGKDEAQRSDMELLRFSLLIQSWLGPVOL 118

```

RESULT 14
SOMA_MOUSE
ID SOMA_MOUSE STANDARD; PRI: 217 AA.
AC Q6680;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Somatotropin precursor (Growth hormone).
GN GH1.
OS Nycticebus pygmaeus (Pygmy slow loris).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Strepsirrhini; Lorisae; Nycticebus.
OX NCBI_TaxID=101278;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
KA Wallis O.C., Zhang Y.P., Wallis M.;
RT "Cloning and characterisation of the gene encoding slow loris growth
  hormone."
RL Submitted (Aug-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Plays an important role in growth control. Its major
CC role in stimulating body growth is to stimulate the liver and
CC other tissues to secrete IGF-1. It stimulates both the
CC differentiation and proliferation of myoblasts. It also stimulates
CC amino acid uptake and protein synthesis in muscle and other
CC tissues.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
EMBL: AJ237562; CAC03504.1;
DR HSSP: P01246; 1B8T.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; Hormone; 1.
DR PRINTS: PR00836; SOMATOTROPIN.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW Hormone; Pituitary; Signal.
FT SIGNAL 1 27 BY SIMILARITY.
FT CHAIN 28 217 SOMATOTROPIN.
FT DISULFID 79 190 BY SIMILARITY.
FT DISULFID 207 215 BY SIMILARITY.
SQ SEQUENCE 217 AA; 24595 MW; 7F590E77E5D0065F5 CRC64.

Query Match 38.5%; Score 305.5; DB 1; Length 217;
Best Local Similarity 65.2%; Pred. No. 9.2e-23;
Matches 60; Conservative 14; Mismatches 17; Indels 1; Gaps 1;

OY 2 FPTIPLSLFDNAVLRAERHQLAFTDYQEFEEAYIPKEQKYSYHQLQNTQSLSESSIPT 61
DB 28 FPMPLSSLSFANAVLRAGHLHQLAFTDYQEFEEAYIPKEQKYSYHQLQNTQSLSESSIPT 61
OY 62 PSNREETOQKSNLELLRISLLIQLSWLPVQL 93
DB 67 PTGKDEAQRSDMELLRFSLILLIQLSWLPVQL 118

RESULT 15
SOMA_MOUSE
ID SOMA_MOUSE STANDARD; PRI: 216 AA.
AC P06880;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Somatotropin precursor (Growth hormone).
GN GH1 OR GH.

```

```

Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=85261158; PubMed=2951252;
KA Linzer D.I.H., Talamantes F.;
RT "Nucleotide sequence of mouse prolactin and growth hormone mRNAs and
  expression of these mRNAs during pregnancy."
RL J. Biol. Chem. 260:9574-9579(1985).
RN [2]
RP SEQUENCE FROM N.A.
RC SIRAN-FZTDJ; TISSUE=Liver;
KA Das P., Meyer L., Seyfert H.-M., Brockmann G., Schwerin M.;
RT "Structure of the growth hormone-encoding gene and its promoter in
  mice."
RL Gene 169:209-213(1996).
CC -!- FUNCTION: Plays an important role in growth control. Its major
CC role in stimulating body growth is to stimulate the liver and
CC other tissues to secrete IGF-1. It stimulates both the
CC differentiation and proliferation of myoblasts. It also stimulates
CC amino acid uptake and protein synthesis in muscle and other
CC tissues.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
EMBL: X02891; CAA26650.1;
DR EMBL: Z46663; CAA8558.1;
DR PIR: B23911; STMS.
DR HSSP: P01246; 1B8T.
DR MGI: MGI:95707; Gh.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; Hormone; 1.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW Hormone; Pituitary; Signal.
FT SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 27 216 SOMATOTROPIN.
FT DISULFID 78 189 BY SIMILARITY.
FT DISULFID 206 214 BY SIMILARITY.
SQ SEQUENCE 216 AA; 24715 MW; 98666AFAE25D65FC CRC64.

Query Match 38.2%; Score 304.5; DB 1; Length 216;
Best Local Similarity 64.8%; Pred. No. 9.7e-23;
Matches 59; Conservative 14; Mismatches 17; Indels 1; Gaps 1;

OY 2 FPTIPLSLFDNAVLRAERHQLAFTDYQEFEEAYIPKEQKYSYHQLQNTQSLSESSIPT 61
DB 27 FPMPLSSLSFANAVLRAGHLHQLAFTDYQEFEEAYIPKEQKYSYHQLQNTQSLSESSIPT 61
OY 62 PSNREETOQKSNLELLRISLLIQLSWLPVQ 92
DB 86 PTGKDEAQRSDMELLRFSLILLIQLSWLPVQ 116

Search completed: September 16, 2003, 12:38:52
Job time : 12.6732 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Computer 111

OM protein - protein search, using sw model

Run on: September 15, 2003, 12:34:00 : Search time 36.1468 Seconds
(without alignments)
1069.675 Million cell updates/sec

Title: US-09-423-100-7

Perfect score: 797

Sequence: 1 MEPTPLSLRFNAM-RAHF.....IVDQCINISGIVLYENYN 150

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 820525 seqs, 25807264 residues

Total number of hits satisfying chosen parameters: 830-25

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 10%

Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_humani:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organello:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_protein:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaei:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	436	54.7	217	6	Q8WNE0
2	427.5	52.6	245	4	G14644
3	407.5	51.1	217	6	Q07369
4	396	49.7	212	6	Q07368
5	396	49.7	217	6	Q07367
6	381	47.8	217	6	G14407
7	348	43.7	217	6	Q8WNE0
8	341	42.8	202	4	G14643
9	322.5	40.5	217	6	Q8WNE0
10	306.5	38.5	216	11	Q06615
11	301.5	37.8	216	6	Q8WNE0
12	301.5	37.6	216	6	Q8WNE0
13	298.5	37.5	216	11	Q8WNE0
14	297.5	37.3	204	6	Q8WNE0
15	297.5	37.3	216	11	Q8WNE0
16	297	37.3	217	6	Q8WNE0

17	290.5	36.4	192	6	Q8TU21
18	289.5	36.3	192	6	Q8TUW9
19	287.5	36.1	190	11	Q9JRG0
20	286.5	35.9	178	6	Q95M25
21	285.5	35.8	217	6	Q8BEC0
22	285.5	35.8	217	6	Q8BEC9
23	285	35.8	167	4	P78451
24	283.5	35.6	178	6	Q95M36
25	280.5	35.2	217	6	Q8B957
26	271.5	34.1	110	6	Q8HXV2
27	265.5	33.4	110	11	Q91X13
28	265.5	33.3	143	6	Q95240
29	261.5	32.8	218	13	Q9PU72
30	254	31.9	110	6	Q8WNE0
31	248.5	30.9	145	6	Q9HNR4
32	234	29.4	199	4	Q14406
33	233.5	29.3	195	13	Q91386
34	224.5	28.2	106	13	Q91807
35	201.5	25.3	110	13	Q8A268
36	197	24.7	106	13	Q5DDE5
37	195.5	24.5	106	13	Q5DDE4
38	193.5	24.3	111	13	Q98TB0
39	193.5	24.3	110	13	Q902V1
40	191.5	24.0	111	13	Q98TA7
41	187.5	23.5	67	13	Q98TA9
42	187.5	23.5	108	13	Q98TB1
43	186.5	23.4	93	6	Q8HXV8
44	185.5	23.3	91	13	Q98TB2
45	177.5	22.3	209	13	Q8AXX9

ALIGNMENTS

RESULT 1

Q8WNE0 PRELIMINARY: PRT: 217 AA.
AC Q8WNE0;
DT 01-MAR-2002 (TRENBLrel: 20, Created)
D: 01-MAR-2002 (TRENBLrel: 20, Last sequence update)
DE 01-MAR-2003 (TRENBLrel: 23, Last annotation update)
DE Growth hormone.
GN GH-N.
OS Ateles geoffroyi (Black-handed spider monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Ateleinae; Ateles.
CX NCBI_taxid=9509;
RN 11.
RP SEQUENCE FROM N.A.
RA Kevo, A., Esquivel, D., Santiago D., Bartera-Saldana H.:
RT "Independent duplication of the growth hormone gene in three
RT Anthropoidea lineages".
PL Submitted (APR-2001) to the EMBL/GenBank/DDBP databases.
PR EMBL: AF374234; AAL72286.1; .
RA InterPro: IPR001400; Serratotrocin.
RA Pfam: PF00103; Hormone;
DA PROSITE: PS00256; SOMATOTROPIN_1; 1.
DA PROSITE: PS00338; SOMATOTROPIN_2; 1.
SQ SEQUENCE: 217 AA: 24894 MW: 4258295FF41EFAA85 CRC64;

Query Match 54.7% Score 436; DB 6; Length 217;
Post Local Similarity 66.9%; Pred. No. 1.4e-38;
Matches 97; Conservative 8; Mismatches 24; Indels 16; Gaps 4;

QY 2 FPTPLSLRFNAM-RAHF.....IVDQCINISGIVLYENYN 150
ED 27 FPTPLSLRFNAM-RAHF.....IVDQCINISGIVLYENYN 150
QY 62 PSNRETOCKNLELLRLISLLIOSWLPFVQLGTGPRVNOHLCG-----SHLVE 111
DB 47 PASKRETOCKNLELLRLISLLIOSWLPFVQLGTGPRVNOHLCG-----SHLVE 111
QY 112 ALYIVCG--ERGFFYTPKTRGIVEQ 134


```

OS Macaca mulatta (Rhesus Macaque).
OC Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi:
OC Mammalia: Eutheria: Primates: Catarrhini: Corropithecoidea:
OC Cercopithecoidea: Macaca.
OC NCBI_TaxID=9544;
CX
RN [1]
RP
RC SEQUENCE FROM N.A.
RC TISSUE=Midpregnancy placenta
RX MEDLINE=94008724; PubMed=9404617;
RX GLOS T.G.; Durrain Y., Fisher J.M., Fowler P.J.;
RT *Cloning of four growth hormone/chorionic somatomotropin-related
RT complementary deoxyribonucleic acids differentially expressed during
RT pregnancy in the rhesus monkey placenta.*;
RL Endocrinology 133:1743-1752(1994);
DR EMBL: L16552; AAA1839.12;
DR HSSP: PC1241; IAXI.
DR InterPro: IPR01400; Somatotropin.
DR Pfam: PF00103; hormone_1.
DR PRINTS: PR00836; SOMATOTROPIN.
DR PROSITE: PS00338; SOMATOTROPIN_2.1.
SQ SEQUENCE 217 AA: 24542 MW  FFSAAAR915231F2HC CRC64:
Query Match 49.74; Score 156; 29 6; Length 217;
Best Local Similarity 62.24; Pred. No. 2,76-34;
Matches 74; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 3 PTPLSLRPFNANLRAHLHQLAFTYQSEFAYIPKPKYSLNFTQSLSSSSSLPI 62
DB 28 PSVLSLPLDHAMIAQHLHQLAFTYQSEFAYIPKPKYSLNFTQSLSSSSSLPI 67
QY 63 SNREETQCKSNLELR:SLSLIQSNLEPVG 92
DB 88 SNLEETQCKSNLELR:SLSLIQSNLEPVG 117
RESULT 6
Q14407 PRELIMINARY; PRT: 217 AA.
AC Q14407;
DT 01-NOV-1996 (TRENDArel. 01, created)
DT 01-NOV-1996 (TRENDArel. 01, last sequence update)
DT 01-MAR-2003 (TRENDArel. 23, last annotation update)
DE Chorionic somatomotropin CS-2 (Chorionic somatomotropin hormone
DE 2).
DE Homo sapiens (Human).
OC Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi:
OC Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
OC NCBI_TaxID=9606;
EN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89307277; PubMed=274769;
RA Chen E.Y., Liao Y.C., Smith D.H., Bartera-Saldana H.A., Gehlens A.F.,
RA Seeburg P.H.;
RT "The human growth hormone locus: nucleotide sequence, biology, and
RT evolution.";
RL Genomics 4:479-497(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91102558; PubMed=1980152;
RA Vencak-Jones C.L., Phillips J.A.;
RT "Hot spots for growth hormone gene deletions in homoious regions
RT outside of Alu repeats.";
RL Science 250:1745-1748(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBSJ databases.
EN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBSJ databases.

```

```

OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo;
OX NCBI_TaxID=9606;
RN 1;
RP SEQUENCE FROM N.A.
RX TISSUE-Term placenta;
RX MEDLINE=9437377; PubMed=5709943;
RA Boguszewski C.L., Svensson P.A., Jansson T., Clark R.,
RA Carlsson L.M.S., Carlsson B.
RT "Cloning of two novel growth hormone transcripts expressed in human
RT placenta.";
RL J. Clin. Endocrinol. Metab. 83:2878-2885(1998).
DR EMBL: AF006060; AF571828.1;
DR HSSP: P01241; IAZ2.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone_1.
DR PRINTS: PR00836; SOMATOTROPIN.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW Signal.
FT SIGNAL.
SQ SEQUENCE 202 AA; 23128 MW; 38B64D31A9197C6 CRC64;

Query Match 42.8%; Score 341; DB 4; Length 202;
Best Local Similarity 65.8%; Pred. No. 1.9e-28;
Matches 75; Conservative 4; Mismatches 9; Indels 26; Gaps 2;

QY 2 PTPILSRFDNAMLRAHRLHQLAFDTYOEFEAYIPKEQYSLFONPQTSLSSESPT 61
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
27 PTPILSRFDNAMLRAHRLHQLAFDTYOEFEAYIPKEQYSLFONPQTSLSSESPT 71
QY 62 PSNREETOQSKNLELRISLLIQSWLEPVLQ-----GTGPRFVNQHL 104
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
72 PSNREETOQSKNLELRISLLIQSWLEPVLQ-----GTGPRFVNQHL 128

RESULT 9
Q8MI74 PRELIMINARY; PRT; 217 AA.
ID Q8MI74
AC Q8MI74
DT 01-OCT-2002 (TRENHRel. 22, Created)
DT 01-OCT-2002 (TRENHRel. 22, Last sequence update)
DT 01-MAR-2003 (TRENHRel. 23, Last annotation update)
DE Growth hormone-like protein precursor.
GN GH26.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Callitrix;
OX NCBI_TaxID=9483;
RN 1;
RP SEQUENCE FROM N.A.
RX Wallis O.C., Wallis M.;
RT "Characterisation of the GH gene cluster in a New-World monkey, the
RT marmoset (Callithrix jacchus).";
RL J. Mol. Endocrinol. 9:0-0(2002).
DR EMBL: AJ489811; CA634012.1;
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone_1.
DR PRINTS: PR00836; SOMATOTROPIN.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW Signal.
FT SIGNAL.
FT CHAIN 1 26 POTENTIAL.
FT CHAIN 27 217 GROWTH HORMONE-LIKE PROTEIN 6.
SQ SEQUENCE 217 AA; 25177 MW; 5ECF148798278F1A CRC64;

Query Match 40.5%; Score 322.5; DB 6; Length 217;
Best Local Similarity 64.7%; Pred. No. 2e-26;
Matches 66; Conservative 13; Mismatches 22; Indels 1; Gaps 1;

QY 3 PTPILSRFDNAMLRAHRLHQLAFDTYOEFEAYIPKEQYSLFONPQTSLSSESPT 62
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
28 PTPILSRFDNAMLRAHRLHQLAFDTYOEFEAYIPKEQYSLFONPQTSLSSESPT 87

```

```

QY 63 SNKREETOQSKNLELRISLLIQSWLEPVLQGTGPRFVNQHL 104
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
88 FRKEEMLGKSNVLAHLISLLIQSWLEPVLQGLGSIPIANSOL 128

RESULT 10
Q70615 PRELIMINARY; PRT; 216 AA.
ID Q70615
AC Q70615
DT 01-AUG-1998 (TRENHRel. 07, Created)
DT 01-AUG-1998 (TRENHRel. 07, Last sequence update)
DT 01-MAR-2003 (TRENHRel. 23, Last annotation update)
DE Growth hormone precursor.
OS Spalax leucodon ehrenbergi (Ehrenberg's mole rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Spalacinae;
OC Nannospalax
OX NCBI_TaxID=30637;
RN 1;
RP SEQUENCE FROM N.A.
RX MEDLINE=99124645; PubMed=9924177;
RA Lioupis A., Nevo E., Wallis M.;
RT "Cloning and characterisation of the gene encoding mole rat (Spalax
RT ehrenbergi) growth hormone.";
RL J. Mol. Endocrinol. 22:29-36(1999).
DR EMBL: AF005819; CA006716.1;
DR HSSP: P01241; IAX1.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone_1.
DR PRINTS: PR00836; SOMATOTROPIN.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW Signal.
FT SIGNAL.
FT CHAIN 1 26 POTENTIAL.
FT CHAIN 27 216 GROWTH HORMONE.
SQ SEQUENCE 216 AA; 24627 MW; EA8A8A523BACADFE CRC64;

Query Match 39.5%; Score 306.5; DB 11; Length 216;
Best Local Similarity 65.9%; Pred. No. 1e-24;
Matches 60; Conservative 13; Mismatches 17; Indels 1; Gaps 1;

QY 2 PTPILSRFDNAMLRAHRLHQLAFDTYOEFEAYIPKEQYSLFONPQTSLSSESPT 61
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
27 PTPILSRFDNAMLRAHRLHQLAFDTYOEFEAYIPKEQYSLFONPQTSLSSESPT 85
QY 62 PSNREETOQSKNLELRISLLIQSWLEPVLQ 92
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
86 PTGKEEACGRSDMEJLRSLLIQSWLGVLQ 116

RESULT 11
Q8MI73 PRELIMINARY; PRT; 216 AA.
ID Q8MI73
AC Q8MI73
DT 01-OCT-2002 (TRENHRel. 22, Created)
DT 01-OCT-2002 (TRENHRel. 22, Last sequence update)
DT 01-MAR-2003 (TRENHRel. 23, Last annotation update)
DE Growth hormone precursor.
GN GH.
OS Delphinus delphis (Saddleback dolphin) (Black sea dolphin).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti; Delphinidae;
OC Delphinus.
OX NCBI_TaxID=9728;
RN 1;
RP SEQUENCE FROM N.A.
RX FISSUE-Silver;
RA Manion Z., Wallis O.C., Wallis M.;
RT "Cloning and characterisation of the GH gene from the common dolphin
RT (Delphinus delphis).";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ492191; CAD37292.1;

```


<hr/>							
SEQUENCE FROM N.A.							
KA	Nayana T.N.K., Lieberhaber S.A.; "Mouse Growth Hormone Locus: Nucleotide Sequence and Phylogenetic Analyses."						
KT	Submitted (AUG-1995) to EMBL/GenBank/CDBJ databases.						
KI	Submitters:						
DR	EMBL; U34362; AAC99988.1;						
FT	HSSP; P01241; IAXI.						
DP	SIGNAL	1	26	POTENTIAL			
FT CHAIN	27	216	GROWTH HORMONE				
SEQ SEQUENCE	216 AA; 24505 KW; IE467A3ACQFES02 CRC64;						
Query Match 37.8%; Score 301.5; DB 6; Length 216; Best Local Similarity 64.8%; Pred. No. 3.4e+24; Matches 59; Conservative 14; Mismatches 17; Indels 1; Gaps 1;							
OY	2 FETPLSRFNAMLRHRLHOLAADTYOEFEAYVPEEAEYIKPKYSFLQNPTSLFSSESIPT 61 II : IIII : II : IIII : III : IIII : III : IIII : III :						
DG	27 FRAMSLSLFANAVLAHQHLHOLAATYKEERAYVEGGKYS-IQTQAFCFSETIPA 85 II : III : IIII : III : IIII : III : IIII : III : III :						
OY	62 PSNREETOXSNLEILISLLIQSWLPVQ 92 II : III : IIII : III : IIII : IIII : IIII : IIII :						
DG	86 PTGKDFAQRSVDLLRFRSLLLIQSGLGVQ 116 II : III : IIII : IIII : IIII : IIII : IIII : IIII :						
RESULT 12							
O8HYE5 PRELIMINARY: PRD: 216 AA.							
ID O8HYE5 AC Q8HYE5							
DI 01-MAR-2003 (TREMBLrel_23, Created)							
DT 01-MAR-2003 (TREMBLrel_23, Last sequence update)							
DE 01-MAR-2003 (TREMBLrel_23, Last annotation update)							
OS Growth hormone precursor.							
OC Alluropoda melanoleuca (GIant panda).							
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
OC Mammalia; Eutheria; Carnivora; Fissipedia; Ursidae; Artiodactyla;							
CN NCBI_TaxID=9645;							
RN [1]							
RP SEQUENCE FROM N.A.							
RC ISSUES-Pituitary;							
RA Liao M., Zhu M., Zhang A.;							
RT "Cloning and expression of cDNA encoding growth hormone from							
RT Alluropoda melanoleuca.";							
RL Submitted (AUG-2002) to the EMBL/Genbank/CDBJ databases.							
DR EMBL; AF545636; AN77228.1;							
KW Signal.							
FT SIGNAL	1	26	POTENTIAL				
SEQ SEQUENCE	216 AA; 24283 KW; 44EC17EC4BGRG56 CRC64;						
Query Match 37.8%; Score 301.5; DB 6; Length 216; Best Local Similarity 64.8%; Pred. No. 3.4e+24; Matches 59; Conservative 14; Mismatches 17; Indels 1; Gaps 1;							
OY	2 FETPLSRFNAMLRHRLHOLAADTYOEFEAYVPEEAEYIKPKYSFLQNPTSLFSSESIPT 61 II : IIII : II : IIII : III : IIII : III : IIII : III :						
DG	27 FRAMSLSLFANAVLAHQHLHOLAATYKEERAYVEGGKYS-IQTQAFCFSETIPA 85 II : III : IIII : III : IIII : III : IIII : III : III :						
OY	62 PSNREETOXSNLEILISLLIQSWLPVQ 92 II : III : IIII : III : IIII : IIII : IIII : IIII :						
DG	86 FTGKDFAQRSVDLLRFRSLLLIQSGLGVQ 116 II : III : IIII : IIII : IIII : IIII : IIII : IIII :						
RESULT 13							
O9R2C3 PRELIMINARY: PRD: 216 AA.							
ID O9R2C3 AC Q9R2C3							
DI 01-MAY-2000 (TREMBLrel_13, Created)							
DT 01-MAY-2000 (TREMBLrel_13, Last sequence update)							
DE 01-MAR-2003 (TREMBLrel_23, Last annotation update)							
OS Mus musculus (Mouse).							
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.							
CN NCBI_TaxID=10090;							
RN [1]							
<hr/>							
SEQUENCE FROM N.A.							
KA	Nayana T.N.K., Lieberhaber S.A.; "Mouse Growth Hormone Locus: Nucleotide Sequence and Phylogenetic Analyses."						
KT	Submitted (AUG-1995) to EMBL/GenBank/CDBJ databases.						
KI	Submitters:						
DR	EMBL; U34362; AAC99988.1;						
FT	HSSP; P01241; IAXI.						
DP	SIGNAL	1	26	POTENTIAL			
FT CHAIN	27	216	GROWTH HORMONE				
SEQ SEQUENCE	216 AA; 24505 KW; IE467A3ACQFES02 CRC64;						
Query Match 37.8%; Score 301.5; DB 6; Length 216; Best Local Similarity 64.8%; Pred. No. 3.4e+24; Matches 59; Conservative 14; Mismatches 17; Indels 1; Gaps 1;							
OY	2 FETPLSRFNAMLRHRLHOLAADTYOEFEAYVPEEAEYIKPKYSFLQNPTSLFSSESIPT 61 II : IIII : II : IIII : III : IIII : III : IIII : III :						
DG	27 FRAMSLSLFANAVLAHQHLHOLAATYKEERAYVEGGKYS-IQTQAFCFSETIPA 85 II : III : IIII : III : IIII : III : IIII : III : III :						
OY	62 PSNREETOXSNLEILISLLIQSWLPVQ 92 II : III : IIII : III : IIII : IIII : IIII : IIII :						
DG	86 PTGKDFAQRSVDLLRFRSLLLIQSGLGVQ 116 II : III : IIII : IIII : IIII : IIII : IIII : IIII :						
RESULT 15							
O9JKM4 PRELIMINARY: PRD: 204 AA.							

[illegible]

```

ID Q9JAM4 PRELIMINARY: PRT: 216 AA.
AC Q9JAM4:
DI 01-OCT-2003 (Tremblrel: 15, (recreated)
DI 01-OCT-2000 (Tremblrel: 15, last sequence update)
DI 01-MAR-2003 (Tremblrel: 23, last annotation update)
DE Growth hormone precursor.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriognathia; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
SEQUENCE FROM N.A.
RA Odorico D.M., Fuller P.C., Herington A.C.:
RT "Cloning and sequence of quinea pig growth hormone (GH).";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DS EMBL: AF233653; AAF36409.1;
DS HSP: P01241; IAXI.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00103; hormone; 1.
DR PRINTS: PRO0836; SOMATOTROPIN.
DR PROSITE: PSC0266; SOMATOTROPIN_1; 1.
DR PROSITE: PSC0338; SOMATOTROPIN_2; 1.
KW Signal.
FT SIGNAL 1 26 POTENTIAL..
FT CHAIN 27 215 GROWTH HORMONE.
SQ SEQUENCE 216 AA: 2482 MW: 4596BEL19H08CD3 CRC64:
Query Match 37.3%; Score 297.5; FR 11; Length 216;
Best Local Similarity 63.7%; Pred. No. 9, 20-24;
Matches 58; Conservative 14; Mismatches 18; Indels 1; Gaps 1;
QY 2 FTTPISRLFDNMLRAHRLHQIADTYQEPFEAYIPKCKYSLQNFQTSLSFSESFT 40
DB 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
27 PPAMPLSSFGNAVLRAQHLHQIAIYKEPERTYIPESQKYSIHNTQTAFQFSELTIA 65
QY 62 PSNRETCCKSNLRLRLSLCZQSWLHPVQ 92
DB 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
86 PDKKFAQRSDVELLHPSLLIQSWLGIHQ 116

```

Search completed: September 16, 2003, 12:40:01
 Job time : 37.1868 secs